Author:
Tromans, Robert

Title:
Urea Based Synthetic Lectins

A Biomimetic Receptor for Glucose

Urea Based Synthetic Lectins: A Biomimetic Receptor for Glucose

Robert Tromans

Supervisors:

Prof. A. P. Davis, Prof. M. P. Crump

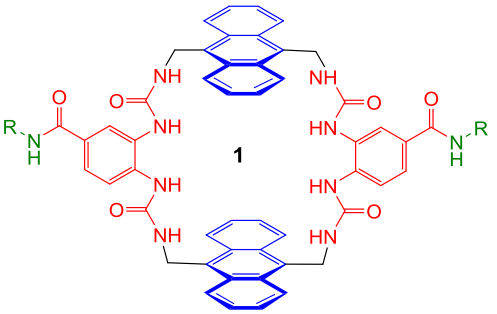
July 2018

A thesis submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science.

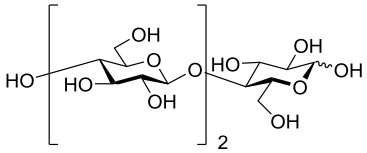
Word Count: 56,671

Abstract

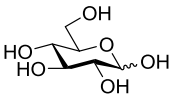
The binding of polar organic molecules in water is one of the most difficult problems in molecular recognition. Carbohydrates are especially challenging targets, given their complex, hydromimetic and subtly variable structures. As a result, carbohydrate-binding proteins such as lectins often display low affinities and moderate selectivities, especially when compared to other classes of small molecule binding proteins. 'Synthetic lectin' mimics have been reported for several years, but have typically been still weaker, with previous approaches generally consisting of macrocycles employing aromatic surfaces bridged with amide bearing spacer units.



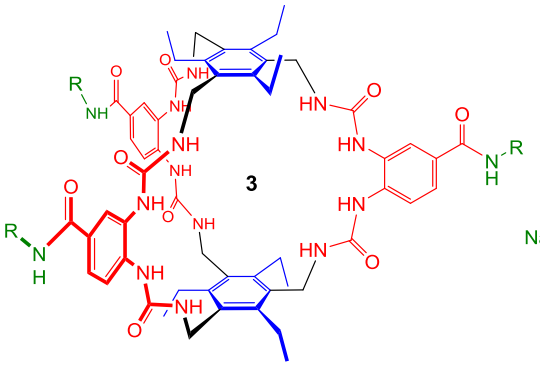
1



2

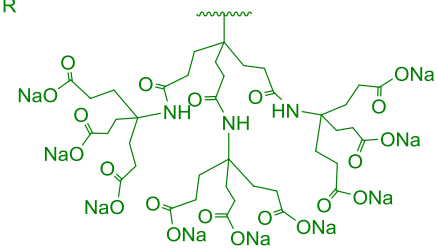


4



3

R = water solubilising group:



This project aimed to investigate a new type of synthetic lectin employing bis-urea spacer units. Incorporation of this urea spacer design into anthracene tetraurea **1** afforded good affinities and selectivities, comparable to natural lectins, for cellodextrin oligosaccharide substrates (such as D-cellotriose **2**). The potential of the urea spacer was fully realised when incorporated into triethylbenzene hexaurea **3** however, which shows binding affinities and selectivities for D-glucose **4** that far exceed any previous synthetic system. Receptor **3** even continues to function in complex biological media (such as human blood serum), enabling preliminary proof of concept for **3** acting as a D-glucose **4** sensor. The performance of **3** even surpasses most natural systems (such as lectins) and therefore can be considered the first true example of a synthetic biomimetic receptor for D-glucose.

Acknowledgements

The author would like to give thanks to Professor Anthony Davis for the guidance and support received throughout the project (especially during the more difficult parts), as well as providing the freedom to work independently and learn through my own discovery. Additional thanks are also given to Professor Matt Crump for assistance with NMR spectroscopy.

Thanks are necessary for the Davis group (past and present): Chris Dias, Alberto Avila Castro, Miriam Wilson, Claire Webster, Sergio Benitez Benitez, Patrick Stewart, Tom Carter, Ash Griffith, Tiddo Mooibroek, Hennie Valkenier, Amir Shermila Kaspar, Charley Renney, Sophie Beattie, James Cooper, Bahar Shirinfar, Nisar Ahmed, Germinal Magro, Pablo Ríos, David Madariaga, Ondřej Jurček, Micke Lisbjerg, Conor Lanphere, Joa Zhao, Vito Zhao, Martin Doll, Eva González Freire, Eléa Wallisky, Kristian Jones, Jack Maynard, Luca Poli, Kelly Chu, Martin Maclaughlin, William Raimbach and all other visitors. Thanks for making the lab such a fun place to be, providing support when the chemistry didn't quite go my way and how to solve those problems, not forgetting the sometimes questionable music choices and all those Friday nights in the pub.

Thank you to the 2013 CDT cohort: Chris Dias, Matt Everett, Rosie Falconer, Stephen Hill, Charlie Jarrett-Wilkins, Nick Lees, Thomas Nunns, Dan O'Flynn, Katy Pellow, Guto Rhys, James Smith, Will Yu and honorary member: Iain Kendall. Thanks for the support, fun times and resilience during my presentations of failed synthesis in the early years. Thanks to Friday lunch bunch for all the falafel.

Thanks to Ziylo Ltd., most notably Laurent Chabanne, Andy Chapman, Mike Orchard, Johnathan Matlock, Michael Tomsett and Harry Destecroix, for helping me broaden my understanding beyond my project and realise the bigger picture, as well as for the additional funding.

Thanks are also necessary for the Bristol Chemical Synthesis Centre for Doctoral Training (BCS-CDT): Kevin Booker-Milburn, Emma Rose, Laura Chavda and the rest of the CDT management for providing the experience and opportunity. Thanks are also necessary for the EPSRC for the funded studentship.

Additional thanks to Miriam Wilson and Andy Chapman for proofreading.

And of course, thanks to my family: Rita, Terry, Katherine and Jenny for all their love and support.

Author’s Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE:.....

Contents

Table of Figures

Table of Tables

Abbreviations

ABCN	1,1'-Azobis(cyclohexanecarbonitrile)
AFM	Atomic force microscopy
Boc	*tert*-butyloxycarbonyl
CD	Circular dichroism
CDI	Carbonyldiimidazole
COSY	homonuclear correlation spectroscopy
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCC	N,N'-Dicyclohexylcarbodiimide
DCM	Dichloromethane
DIPEA	N,N-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMEM	Dulbecco's Modified Eagle Medium
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
ESI	Electrospray Ionisation
EtOAc	Ethyl acetate
Fmoc	Fluorenylmethyloxycarbonyl
FRET	Förster Resonance Energy Transfer
G	Guest
G1L	1st generation “large” dendrimer
G1M	1st generation “medium” dendrimer
G2MM	2nd generation “medium medium” dendrimer
H	Host

HBTU	N,N,N’,N’-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate
HG	Host-Guest complex
HMBC	Heteronuclear Multiple Bond Correlation
HMDS	Hexamethyldisilazane
HOBt	Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear Single-Quantum Correlation
ICD	Induced circular dichroism
IR	Infrared
ITC	Isothermal titration calorimetry
K_a	Association constant
K_d	Dissociation constant
k_{ex}	rate of exchange
L-15	Leibovitz's L-15 Medium
LRMS	Low resolution mass spectrometry
MALDI	Matrix Assisted Laser Desorption/Ionization
MCMM	Monte Carlo molecular mechanics
MRI	Magnetic resonance imaging
MS	Mass spectrometry
MWCO	Molecular weight cut off
NBS	N-Bromosuccinimide
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy

PBS	Phosphate buffered saline
Ppm	Parts per million
ROESY	Rotating frame nuclear Overhauser effect spectroscopy
TEB	1,3,5-triethylbenzene
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TOCSY	Total correlation spectroscopy
UV	Ultraviolet
UV/vis	Ultraviolet-visible spectroscopy

Chapter 1 - Introduction

1.1 Carbohydrates and their recognition by natural and synthetic lectins

The class of organic molecules known as carbohydrates are ubiquitous throughout nature (Figure 1). They provide the main source of energy for many organisms (D-glucose **4**), act as key components in biological systems such as DNA (2-deoxy-D-ribose) and deliver macroscopic structural rigidity when incorporated into biopolymers (D-cellulose **8**and D-chitin **9**).[1,2] In addition to these classical roles, saccharide chains of varying lengths can contain an extraordinary amount of information for biological systems. Combining the subtle stereochemical variance of monosaccharides with possible substitution at any of the hydroxyls, differing rings sizes, glycosidic linkage positions and branching of saccharide chains results in a wealth of potential structural diversity. As a result of this a sequence of six monosaccharides produces more than 10^{12} different potential oligosaccharides. Comparatively six nucleotides generates only 4096 nucleotide sequences and six amino acids produces 6×10^7 possible combinations of polypeptide.[3,4] It should be of no surprise then that the recognition of characteristic sugar chains is instrumental in numerous biological processes, such as cell to cell recognition, protein transport and infection by pathogens. Such processes are mediated by proteins known as lectins that can selectively recognise and bind carbohydrates in a non-covalent manner.[5,6]

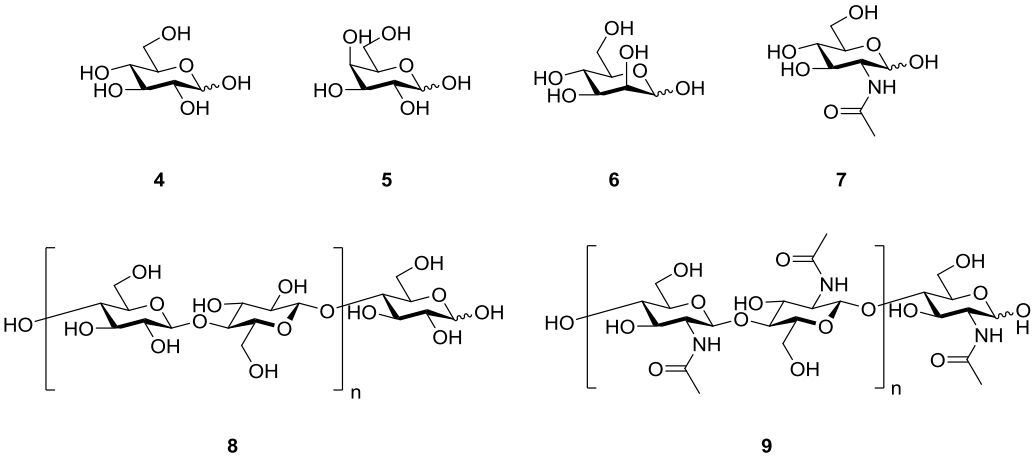


Figure 1 Examples of carbohydrates commonly found in nature: D-glucose **4**, D-galactose **5**, D-mannose **6**, *N*-acetyl-D-glucosamine **7**, D-cellulose **8** and D-chitin **9**.

The affinities and selectivity achieved by natural lectins is impressive when considering the subtle structural variance of carbohydrates. The single stereochemical difference seen for D-glucose **4**, D-galactose **5** and D-mannose **6** results in completely different functions in biology and yet, are very similar in the overall structure of the molecule. The other key consideration is the aqueous medium that biology operates in. Carbohydrates have an outer surface of hydroxyl groups that effectively camouflage the core carbon ring from the polar solvent. This results in most carbohydrates being

hydromimetic (resembling water), thus making the distinction between the substrate and the solvent by the lectin all the more challenging (Figure 2).[7,8]

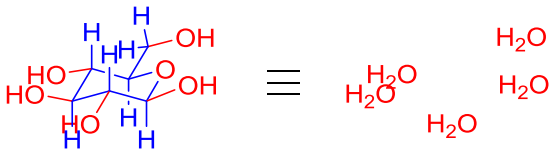


Figure 2 Hydromimetic nature of β-D-glucose, where the outer hydroxyls resemble water.

This high degree of selectivity observed in lectins is achieved through highly complementary binding sites, each specifically tailored to a particular carbohydrate. Multiple chemical interactions are present in the binding site. Hydrogen bonding between the sugar hydroxyls and appropriately oriented polar residues (such as arginine and aspartic acid) are used in conjunction with CH-π interactions between the aliphatic parts of the carbohydrate and apolar aromatic residues (such as tyrosine and phenylalanine). Electrostatic and charge-dipole interactions having been observed in the binding sites of some proteins as well. Lectins typically feature shallow binding pockets and thus the sugar isn't typically fully encapsulated unless bound in a protein dimer (Figure 3), so residues are often orientated to be substrate specific to maximise binding interactions.[9,10]

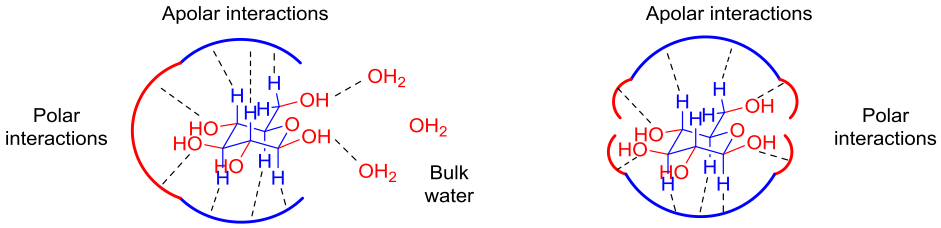


Figure 3 Schematic of simple lectin binding sites for glucose, displaying the general orientation of chemical interactions: apolar interactions (usually CH-π, blue) and polar hydrogen bonding (red). Two types of binding site are also shown, firstly partial encapsulation (left) and full encapsulation by protein dimerisation (right).

While good substrate selectivity is known for natural lectins, relatively weak affinities are also observed with association constants (K_a) of 10^3-10^4 M^{-1} being typical.[11,12] This highlights the inherent challenges of binding carbohydrates in water (structural diversity and solvent competitiveness), and nature sometimes compensates for this through modification of the sugar *via* phosphorylation or chelation with metals (such as manganese or calcium, as seen for C-type lectins).[9,13] It is these moderate affinities however that provides an opportunity to outperform nature through design and synthesis of a 'synthetic lectin'. However, replicating the highly defined binding sites of natural lectins and the innate difficulty of binding carbohydrates in water poses significant challenges. A synthetically accessible host molecule that is capable of selectively binding a target carbohydrate with good affinity would be of

substantial interest; with a host of applications possible such as blood glucose monitoring for diabetes, tumour recognition and potentially new approaches for drug delivery.[14,15]

1.2 Association constants and rates of exchange

The complexation of a guest (i.e. a carbohydrate) to a certain receptor is governed by the following equilibrium (Equation 1)[16]:

$$H + G \underset{K_d}{\overset{K_a}{\rightleftharpoons}} HG \qquad \text{(Eqn. 1)}$$

The association of a free guest (G) to a free host molecule (H) to give a host-guest complex (HG), and its associated binding affinity, is determined by the association constant (K_a) – the greater this value the more favourable the equilibrium is towards forming the HG complex. This association constant can therefore be expressed in terms of concentrations of [HG], [H] and [G] (Equation 2). The dissociation constant (K_d) is the reciprocal of the association constant, but is more commonly used in biological contexts (Equation 3).[16] The vast majority of the work outlined in this thesis will discuss binding affinities in terms of association constants (K_a).

$$K_a = \frac{[HG]}{[H][G]} \qquad \text{(Eqn. 2)} \qquad K_d = \frac{1}{K_a} = \frac{[H][G]}{[HG]} \qquad \text{(Eqn. 3)}$$

Association constants (K_a) are expressed in units of per molar (M^{-1}), but can be expressed in Gibb's free energy of binding (ΔG) and thus joules per mole ($J\ mol^{-1}$) using equation 4:

$$\Delta G = -RT \ln K_a \qquad \text{(Eqn. 4)}$$

Calculation of association constants is therefore dependent on determining the concentrations of [HG], [H] and [G] when known concentrations of host and guest are mixed together. This is typically ascertained through titration of known concentrations of guest into host (or vice versa), with the concentration of host-guest complex determined at each point in the titration through analysis of the experimental data. A combined mixture of host and guest is typically titrated into host, in order to keep the concentration of host constant while the concentration of guest steadily increases throughout the titration.[16]

The most typically used method for determination of association constants is an NMR spectroscopy titration, as this method can also give critical information about binding kinetics and structural information of the host-guest complex. The exact method used to extract the association constant from experimental NMR titration data is dependent on the exchange rate (k_{ex}) of free guest with guest

bound to the host, which is related to binding affinity (K_a), temperature of the system and several other factors. How fast or slow this rate of exchange is relative to the difference between the frequency (chemical shifts in hertz) of bound and unbound forms of the host ($\Delta\omega = \omega_{HG}-\omega_H$) can result in several different outcomes observed in NMR titration spectra (Figure 4).[17]

If rate of exchange is much faster than the difference in frequency between the bound and unbound species (fast exchange, Equation 5), then a time averaged spectrum is obtained – where separate signals for both host (δ_H) and host-guest (δ_{HG}) coalesce into one signal (δ_{obs}, Equation 6). The mole fractions of each species (Equations 7 and 8) then determine the chemical shift of this averaged signal.[16]

$$Fast\ Exchange, K_{ex} \gg \Delta\omega \qquad \text{(Eqn. 5)}$$

$$\delta_{obs} = \delta_H X_H + \delta_{HG} X_{HG} \qquad \text{(Eqn. 6)}$$

$$X_H = \frac{[H]}{[H]_{total}} \qquad \text{(Eqn. 7)} \qquad X_{HG} = \frac{[HG]}{[H]_{total}} \qquad \text{(Eqn. 8)}$$

This is observed during a NMR titration, where addition of guest to host causes a change in chemical shift ($\Delta\delta_{obs}$) for the host signals. At each point in the titration, the chemical shift of the host-guest complex (δ_{HG}) is a representation of the mole fraction of the host-guest complex (X_{HG}) and thus can be used to calculate the concentration of host-guest complex ([HG], Equations 9, 12 and 13). This can in turn be used to determine the remaining concentrations of free host ([H]) and free guest ([G]) using equations 10 and 11, and inputting these values into Equation 2 can yield an association constant (K_a).[16]

$$\delta_{obs} = \delta_H \frac{[H]}{[H]_{total}} + \delta_{HG} \frac{[HG]}{[H]_{total}} \qquad \text{(Eqn. 9)}$$

$$[H] = [H]_{total} - [HG] \qquad \text{(Eqn. 10)} \qquad [G] = [G]_{total} - [HG] \qquad \text{(Eqn. 11)}$$

$$\delta_{obs} = \delta_H \frac{([H]_{total} - [HG])}{[H]_{total}} + \delta_{HG} \frac{[HG]}{[H]_{total}} \qquad \text{(Eqn. 12)}$$

$$[HG] = \frac{[H]_{total}\,(\delta_{obs} - \delta_H)}{(\delta_{HG} - \delta_H)} \qquad \text{(Eqn. 13)}$$

A very different scenario is observed when a host guest system operates with 'slow exchange' kinetics. This occurs when the rate of exchange between bound and unbound guests is much slower than the difference in frequency between bound and unbound species (Equation 14). This results in observation of two discreet signals for both bound and unbound species of host, with signals relating to the host-guest complex (HG) increasing in intensity as more of the host becomes bound by guest during a titration, until the host is saturated with guest and only signals pertaining to the host-guest complex remain. Determination of association constant (K_a) is therefore determined by integration of host-guest signals (Int_{HG}) relative to those of the free host (Int_H). This ratio can be used to determine the concentration of host-guest complex ([HG], Equation 15), and thus [H] and [G] at each point in the titration. Inputting these values into equation 2 as for fast exchange systems yields the association constant (K_a).[16]

$$Slow\ Exchange, K_{ex} \ll \Delta\omega \qquad \text{(Eqn. 14)}$$

$$[HG] = \left(\frac{Int_{HG}}{Int_{HG} + Int_H}\right) \times [H]_{total} \qquad \text{(Eqn. 15)}$$

However, if the rate of exchange (k_{ex}) is similar or equal to the difference in frequency of bound and unbound host species (Δω), then a combination of the effects from fast and slow exchange spectra are observed – termed 'intermediate exchange'. This ultimately results in significant line broadening of host signals as guest is added, until the host signals end up with very large spectral line widths (sometimes completely unobservable in a spectrum) upon saturation with guest. This usually means that no association constant (K_a) can be determined from the experimental data and an alternative method must be used to ascertain binding affinity.[17]

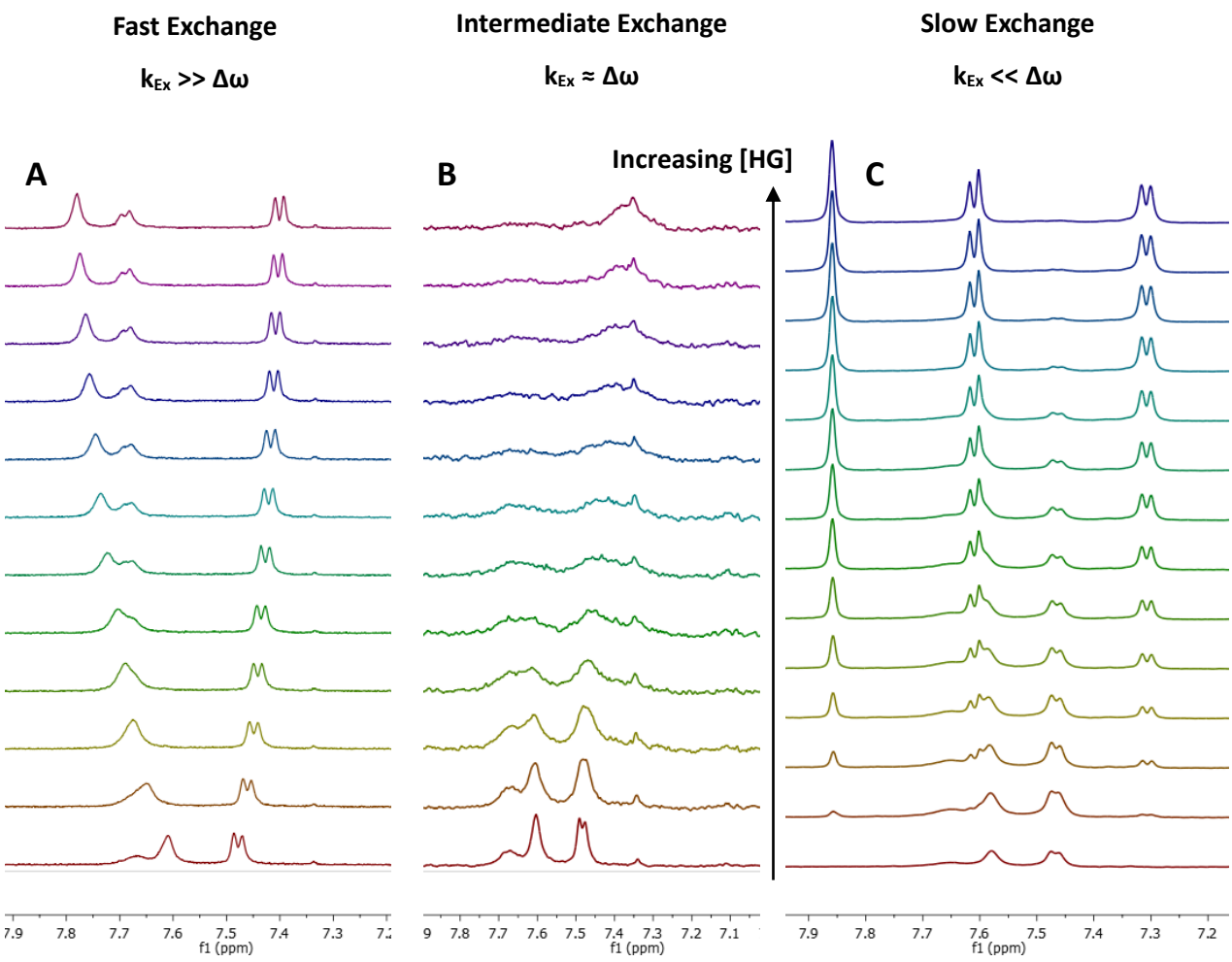


Figure 4 Partial ^{1}H NMR spectra displaying examples of host-guest systems in different exchange rate regimes: fast exchange (A), intermediate exchange (B) and slow exchange (C).

If a secondary method is required to determine binding affinities, theoretically any measurable output that changes upon formation of a host-guest complex can be used. Some common methods used include UV-vis or fluorescence spectroscopy titrations, whereby the change in absorbance or emission intensity (ΔI) is measured as a host molecule becomes complexed to a guest. The change in this 'observable' value (i.e. I_{obs} for UV-vis or fluorescence) can be used to calculate [HG] and thus K_a in much the same way as changes in chemical shift (δ_{obs}) were used in fast exchange systems in NMR titrations.[16]

The majority of the above discussion pertains to a 1:1 binding stoichiometry, where one host binds to one guest. If multiple binding stoichiometries were observed, such as one host binding two guests, then the equations used to calculate [HG] and thus association constant (K_a) would be much more complicated. However, the vast majority of the work discussed in thesis will be related to 1:1 binding stoichiometries. Therefore discussion relating to multiple binding stoichiometries will not covered in great detail in this thesis.

1.3 Carbohydrate recognition by synthetic lectins in organic media

The binding of carbohydrates through non-covalent interactions by a synthetic receptor is more easily attainable in organic solvents than in water, as the competitive nature of the solvent is less of a consideration.[18] The more manageable solubility of the receptor is also of benefit. With no competition from the background solvent, the H-bond interactions between polar groups and the carbohydrate becomes the predominant factor in acquiring good affinities in more apolar media.[19] Thus arrangement of potential H-bond donors must be carefully considered to maximise favourable interactions. The poor solubility demonstrated by some carbohydrates in organic solvents also requires the receptor to encapsulate and essentially solubilise the substrate, although modification to the sugar with aliphatic solubilising groups can circumvent this (such as for octyl glucosides **11** and **12**, Figure 5). While not fully representative of carbohydrate recognition in water, investigations in apolar solvents can allude to potentially active supramolecular architectures that could be developed further to function in an aqueous environment.

One of the earliest examples was reported by Aoyama and co-workers, where calixarene **10** was shown to bind octyl glucosides **11** and **12** *via* a 2:1 host to guest encapsulation in chloroform (Figure 5). Association constants of 3×10^8 M^{-2} were obtained, with the sugar bound by complementary hydrogen bonds between hydroxyl groups of both the substrate and receptor. Future modification to the calixarene scaffold with anionic appendages enabled solubility in aqueous media although a reduction in affinity accompanied this, most likely a result of competitive solvent binding.[20]

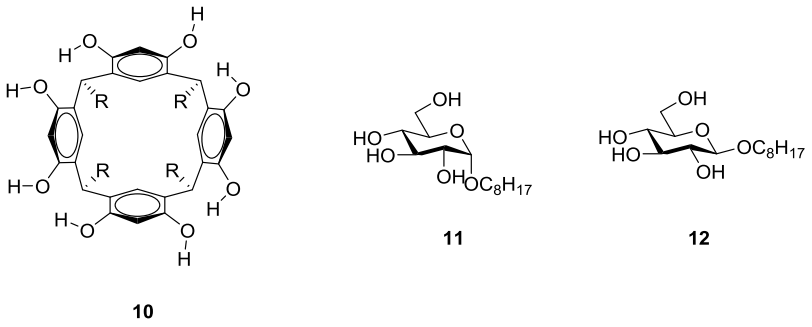


Figure 5 Calixarene **10** synthesised by Aoyama and co-workers showed evidence of binding octyl D-glucosides (α = **11** β = **12**) in chloroform *via* dimerisation of the receptor to encapsulate the sugar (R = $C_{11}H_{23}$).[20]

Another early example was macrocyclic cholaphane **13** reported by the Davis group, which was able to extract methyl β-D-glucoside **14** from water into chloroform (Figure 6). The rigid steroidal backbone, derived from cholic acid, demonstrated a degree of pre-organisation. This conformational inflexibility forced the axial H-bond donor hydroxyls to be orientated towards the substrate binding site, resulting in good affinities and selectivity.[21]

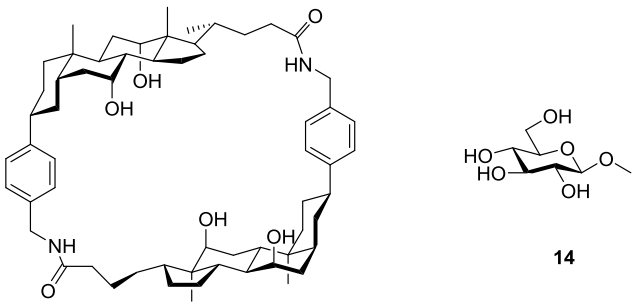


Figure 6 Macrocycle **13** synthesised by Davis and co-workers, which was shown to extract methyl-β-D-glucoside **14** from aqueous solutions in chloroform.[21,]

More recently Roelens *et al* have demonstrated that podand **15** binds octyl β-D-glucoside **12** strongly in both 1:1 and 2:1 receptor/substrate stoichiometries in chloroform (Figure 7). The high degree of pre-organisation is achieved through the 1,3,5-ethyl groups sterically locking the H-bond accepting pyrrole moieties in a tripodal arrangement.[22] This design was then developed further to cryptand **16**, increasing the degree of pre-organisation. This effectively generated a cavity that can fully encapsulate the sugar in a 1:1 binding fashion. This delivered an improved K_a of 4×10^4 M^{-1} for **12** in chloroform, but also increased selectivity to just the β-anomer with no binding observed for the α-anomer.[23]

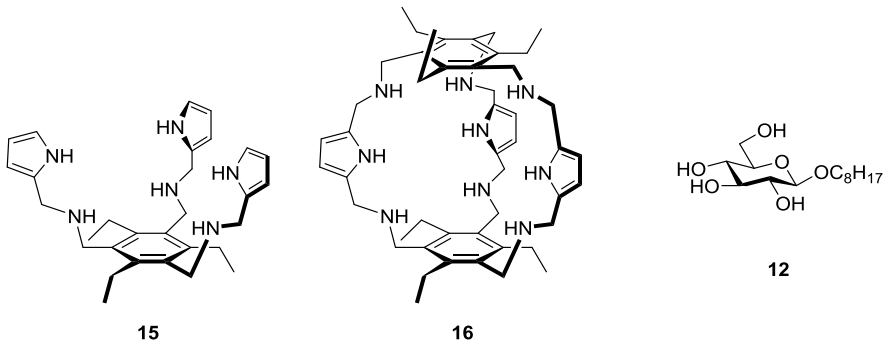


Figure 7 Synthesis and development of podand **12** to cryptand **13** and the resultant increase in pre-organisation by Roelens and co-workers. Both receptors show good affinities for octyl β-D-glucoside **9** in chloroform.[22,23]

The same 1,3,5-triethyl benzene scaffold was also incorporated into a several receptor designs by Mazik and co-workers, with both acyclic and monocyclic variants reported (Figure 8). These also showed appreciable affinities for their corresponding guests, with acyclic receptor **17** affording an affinity of K_a ~1.3×10^5 M^{-1} for **12** in anhydrous chloroform, with two receptors binding to one guest.[24] Evolution of this design to a monocyclic receptor improved affinities, where receptor **18** provided a K_a of ~3×10^5 M^{-1} towards **12** in chloroform.[25] Small additions of water or DMSO severely reduced the binding affinities however, which was speculated to be a result of competitive solvent binding.

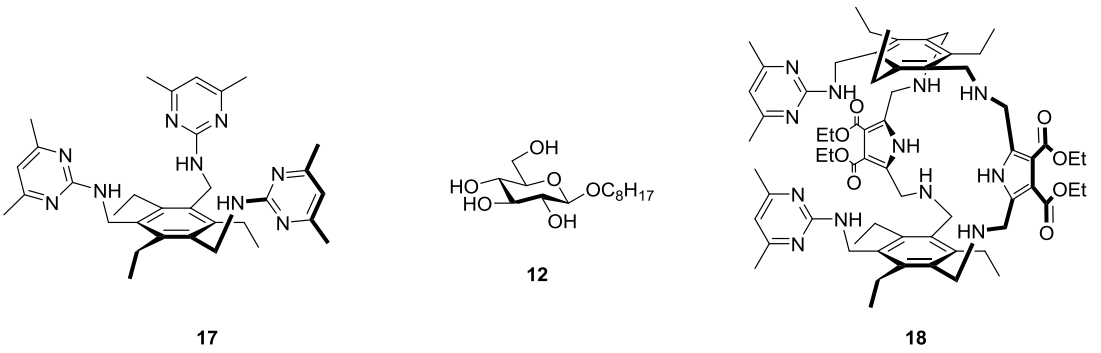


Figure 8 Carbohydrate receptors reported by Mazik *et al*, both of which incorporate the 1,3,5-triethylbenzene scaffold. Acyclic receptor **17** and monocyclic receptor **18** both display affinities for **12** of $K_a > 10^5$ M^{-1} in chloroform.[24,25]

Helical oligoamide **19** was also recently reported by Huc and colleagues to display excellent selectivity and affinities for D-fructose **20** over other common monosaccharides (Figure 9). Association constants of ~3×10^4 M^{-1} were consistently obtained for several variations of the foldamer in a 4:1 mixture of chloroform/DMSO. This activity is a result of the oligoamide elegantly coiling around the substrate in a helical fashion, encapsulating the sugar and delivering both H-bond donors and acceptors where required. Each iteration of the oligoamide was generated *via* repeated modulation of the amide backbone in conjunction with molecular modelling to further refine each design. This process was repeated until the desired cavity size and shape was suitable for the target guest substrate.[26]

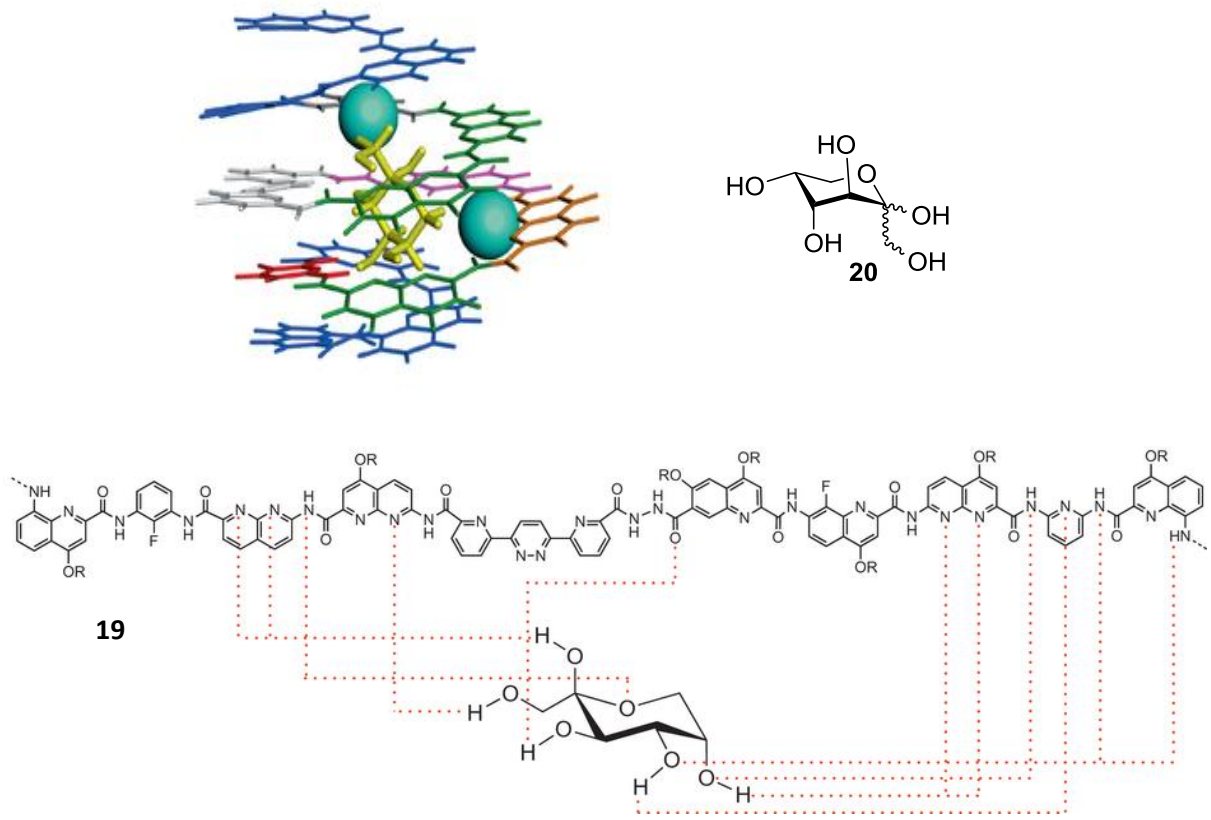
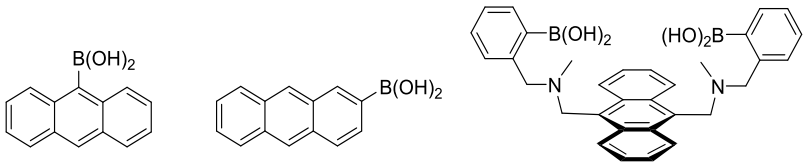


Figure 9 Helical foldamer **19** reported by Huc *et al* shows good selectivity and affinity for D-fructose **20** in 4:1 chloroform/DMSO. The oligoamide helically encapsulates the substrate, providing H-bond donors and acceptors where appropriate. Figure reproduced with permission from Nature Publishing Group.[26]

1.4 Carbohydrate recognition by synthetic lectins in aqueous media

Selectively binding carbohydrates is significantly more difficult in water than in apolar organic solvents. This is due to the competitive association of water molecules in the receptor binding site, which often becomes solvated with a network of water molecules.[18] The receptor must then possess sufficient affinity for a substrate to favourably displace these water molecules into the bulk solution. To achieve this in an aqueous environment means that corresponding interactions to accommodate all regions of the carbohydrate must be incorporated into the binding pocket, as hydrogen bonding is difficult to utilise as the dominant interaction. The other key obstacle is that the receptor must be sufficiently soluble in water and still be able to maintain a defined binding site. Self-binding and aggregation must not occur, as well as the binding site being sufficiently rigid to prevent collapse of the cavity and thus obstruct entry of the substrate. With this in mind, it is unsurprising that few examples exist of synthetic receptors that selectively bind the most common carbohydrates in water through non-covalent interactions.[27]



21 **22** **23**

Figure 10 Anthracenyl boronic acid based receptors **21** and **22** synthesised by Yoon and Czarnik. Receptor **23** was reported by Shinkai and co-workers, and demonstrated improved selectivity for D-glucose as well operation over a broader pH range.[28,29]

While such challenges are apparent for non-covalent recognition systems, they can be mitigated by using a reversible covalent reaction to bind substrates. The most prevalent example to exploit this methodology is the reversible interaction of 1,2 and 1,3-diols with aryl boronic acids.[30] Early examples by Yoon and Czarnik demonstrated that anthracenyl boronic acids **21** and **22** provided a direct fluorescence response upon binding carbohydrates (Figure 10).[28] Selectivity was increased upon modification of these earlier designs by Shinkai *et al* to allow partial encapsulation of the sugar in a shallow binding site, with complementary CH-π interactions provided by the anthracene which also delivered a measurable fluorescence output (receptor **23**, Figure 10).[29] While good affinities for carbohydrates with boronic acids are possible, good selectivities can be difficult to achieve, as boronic acids complex polyols in general.[31] Another consideration is that the equilibrium of the diol co-ordination to boron is highly pH sensitive, generally favouring pH ranges above physiological levels (pH ~9-10), although this can be relieved somewhat by pendant Lewis bases (such as the tertiary amines seen for **23**).[32]

Novel receptor designs based upon non-covalent interactions are rare however. The majority of examples focus on rigid cage-like structures that maintain a centre cavity for the substrate to occupy. These 'temple' style receptors employ complementary apolar and polar interactions, and generally correlate best to all equatorial substrates, such as β-D-glucose (Figure 11), which can be considered as having separate polar and apolar regions (Figure 2). There are two key interactions involved: CH-π interactions between the axial hydrogens of the sugar and the aromatic surfaces of the receptor accompany hydrogen bonds between the equatorial ring of hydroxyls and the polar spacer units. This simplified schematic allows a degree of modularity when designing a receptor and the Davis group has successfully employed this approach towards carbohydrate binding in water.

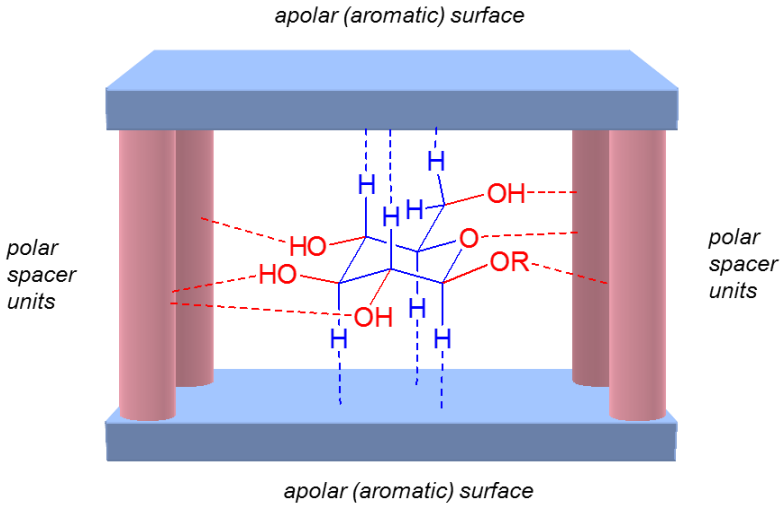


Figure 11 Simple schematic of a 'temple' style receptor designed to bind β-D-glucose. Complementary interactions between apolar (blue) and polar (red) regions of both the substrate and receptor are shown in the binding cavity.

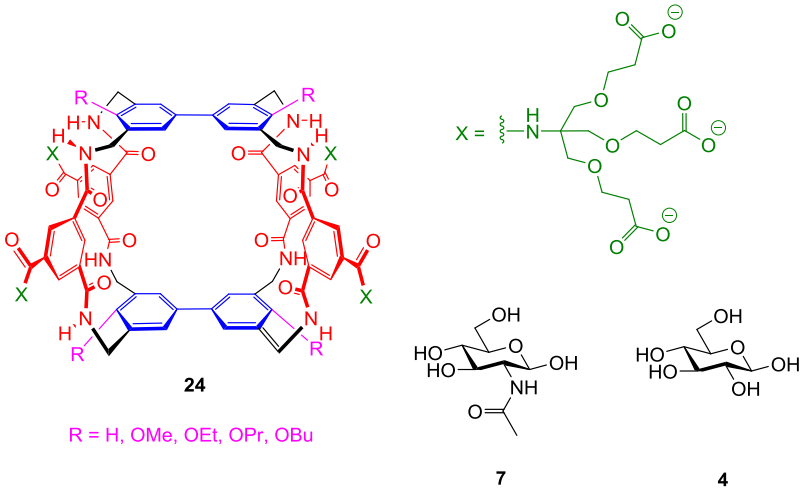


Figure 12 Synthetic lectin **24** reported by the Davis group, showing excellent selectivity for *N*-acetyl-D-glucosamine **7** when R = H.[33] Installation of alkoxy substituents increased affinity for **4** and decreased it for **7**, with the optimum binding observed when R = OPr.[34]

Macrocycle **24** (R = H) was an early attempt in the Davis group to replicate the temple design as a molecular architecture (Figure 12). The biphenyl ceiling and floor is designed to be flush against the sugar to provide CH-π interactions with the axial hydrogens of the substrate. The isophthalamide spacer units provide structural rigidity to maintain the open cavity as well as hydrogen bonding interactions with the equatorial hydroxyls of the carbohydrate. Water solubility was achieved through pendant *tris*-carboxylate chains on each spacer unit, yielding a total of twelve negative charges that make the molecule sufficiently polar to dissolve in aqueous media. This design was found to be very selective for β-*N*-acetylglucosamine **7**, affording a K_a of 630 M^{-1}.[33] This is very comparable to natural

lectin wheat germ agglutinin, which has an affinity for **7** of K_a ~ 730 M^{-1}. Modification of the biphenyl to accommodate terminal alkoxy groups decreased selectivity for **7** but increased selectivity for β-D-glucose **4**, with the propoxy derivative (R = OPr, OC_3H_7) proving the most effective with a K_a of 60 M^{-1}.[34] This is believed to be primarily due to the alkoxy chains pointing inward towards the cavity, thus altering the cavity size and shape.

Evolution of the temple design led to the synthesis of simplified anthracene receptor **25** (Figure 13). This new design was significantly more synthetically accessible than previous iterations and could be accessed after just five steps, which is fifteen less than the synthetic route for **24**. Despite this dramatic simplification there was no loss in performance, with **25** selectively binding β-D-glucose with a K_a of 56 M^{-1} in water.[35] This new receptor opted to substitute out the biphenyl surfaces for anthracene. Biphenyl units have a tendency to twist out of the plane with respect to one another due to steric clash of the *ortho* hydrogens. The rationale with this new design was that the condensed aromatic surfaces would thus provide a more consistent interaction surface for axial hydrogens of the sugar as no such distortion would now be possible.

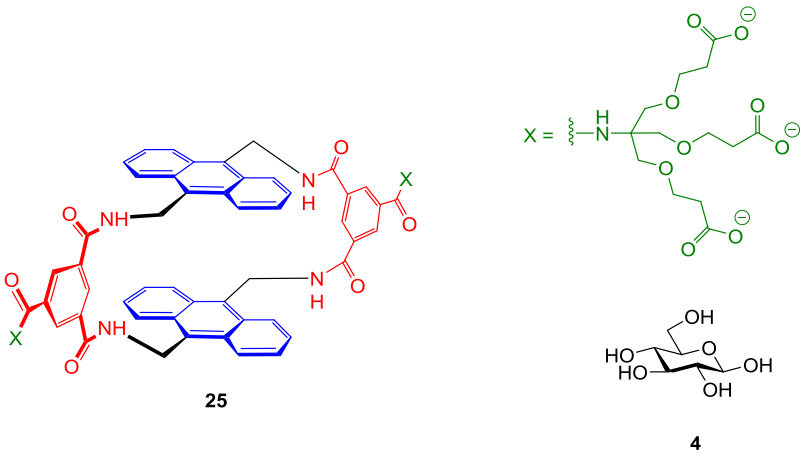


Figure 13 Simplified anthracene based receptor **25**, which shows comparable performance to **24** for binding β-D-glucose **4** in water.[35]

Receptor **25** also halved the number of polar spacer units from four to two, which resulted in a much more flexible molecule. A balance of flexibility and pre-organisation is required for good binding affinities, and it was found that **25** adopted a partially open conformation in solution. Entry of a substrate into the cavity then induces a conformational change which opens up the binding site to fully encapsulate the sugar. Anthracene is inherently fluorescent and this conformational change upon binding induced a fluorescent response that was quantifiable and proportional to the degree of substrate bound within the cavity. This enables a secondary method of sensing glucose binding within the receptor and suggests potential for **25** in the application of glucose monitoring for diabetes.[35]

Extension of the solubilising group to larger dendrimers for **25** increased affinities for β-D-glucose **4**, with 2nd generation "medium-medium" dendritic solubilising group **26** (known within the Davis group as G2MM) providing an improved K_a of 90 M^{-1} in water (Figure 14). A high affinity and selectivity was observed for charged substrates such as D-glucosamine **28**, with **26** giving a binding affinity of K_a ~ 2400 M^{-1}. Extension to higher order dendrimers only appeared to increase this observed affinity and selectivity. Dendrimer **27** affords a K_a of ~7000 M^{-1} for **28** and shows reduced affinity for D-glucose **4**. The affinities observed for **26** and **27** with D-glucosamine **28** are speculated to result from strong charge-charge interactions, which is now possible due to the increased length of the solubilising chains. The terminal carboxylates are now able to enter the cavity and form a salt bridge with the charged groups on the sugar. The larger dendrimers are believed to be under less strain, as well as featuring more carboxylate groups, when forming this salt bridge and thus are capable of higher affinities.[36]

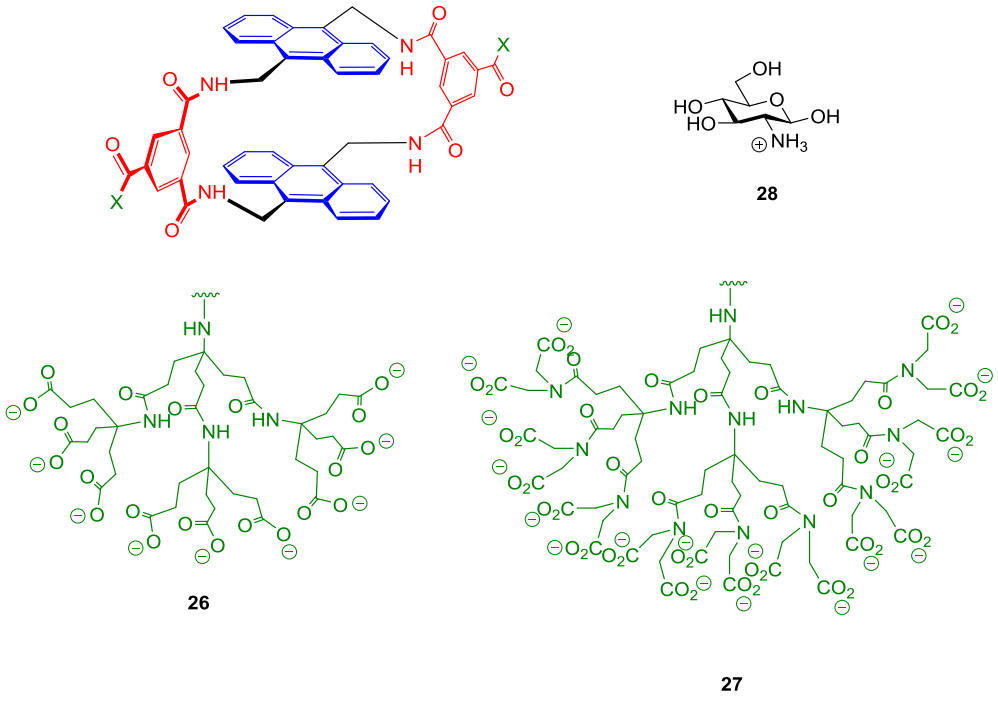


Figure 14 Extension of the dendrimer to G2MM solubilising group **26** and larger dendrimers **27** increases selectivity and affinity for charged sugar D-glucosamine **28** in water.[36]

While receptors **24** and **25** showed good activity for monosaccharides, larger substrates did not display good affinities. Receptors **29** and **30** were based upon **24** but utilised *meta* and *para*-terphenyl surfaces to extend the cavity horizontally, increasing the cavity size to accommodate larger sugars (Figure 15). Indeed, both showed good selectivities for disaccharides but **30** bound substrates with much greater affinity (~5 times stronger binding than for **29**). Receptor **30** afforded a K_a of 3300 M^{-1} for D-cellobiose **31**, which is very comparable to the typical binding strength that natural lectins possess (K_a ~ 10^3-10^4

M^{-1}). Despite this apparent strong binding, further extension of the cavity horizontally to bind larger substrates is not practical and was predicted to result in too much flexibility, making it more difficult to maintain a defined cavity. Therefore such designs are limited to receptors akin to that of **30**.[37,38]

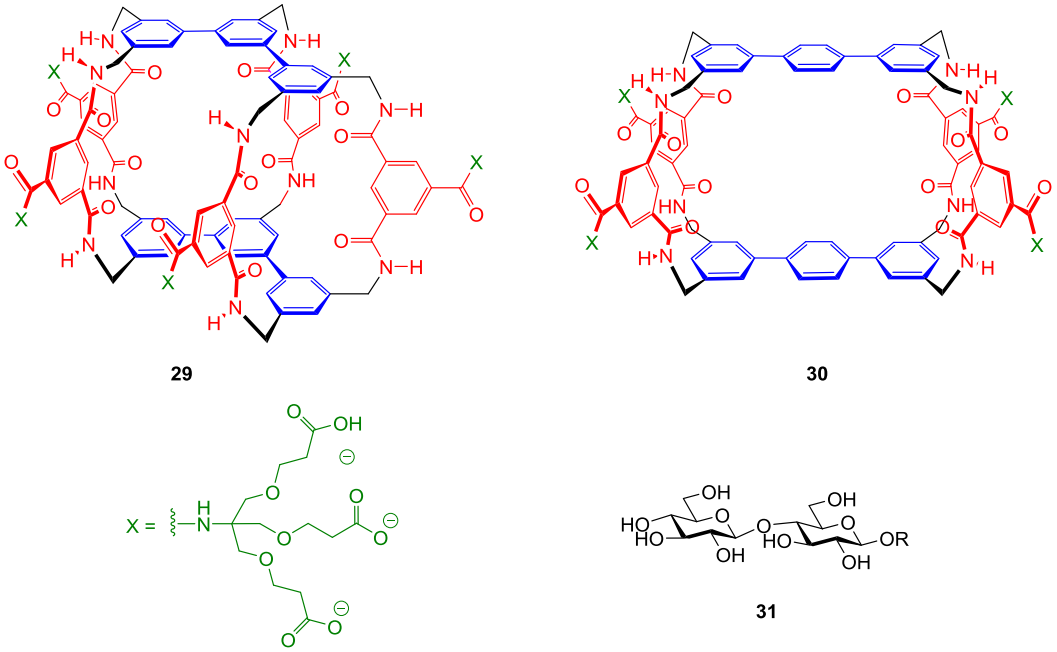
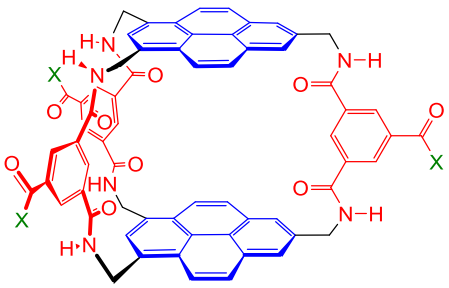


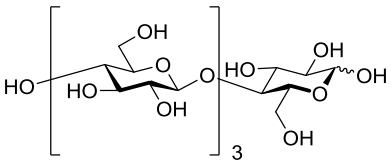
Figure 15 Terphenyl based receptors **29** and **30** feature extended cavities designed to accommodate larger substrates. Receptor **30** shows excellent selectivity and affinity for several key disaccharides including D-cellobiose **31**.[37, 38]

The limitations of substrate size for **29** and **30** were eventually overcome *via* a new method of binding. Pyrene based receptor **32** features a three isophthalamide pillar arrangement to support the open cavity between two pyrene surfaces (Figure 16). While displaying only moderate binding for monosaccharides, the receptor showed strong affinity for D-cellobiose **31** with a K_a of ~5700 M^{-1}; almost a two fold increase over terphenyl receptor **30**. Cellulose derived oligosaccharides of increasing chain length were then subjected to binding strength investigations and it was found that D-cellotetraose **33** had the strongest affinity with a K_a of ~15,000 M^{-1}. Longer oligosaccharides were tested, but a reduction of association constant to K_a ~9,000 M^{-1} was observed and this was speculated to be due to steric interactions between the oligosaccharide and the solubilising groups .[39]



32

X = G2MM solubilising group



33

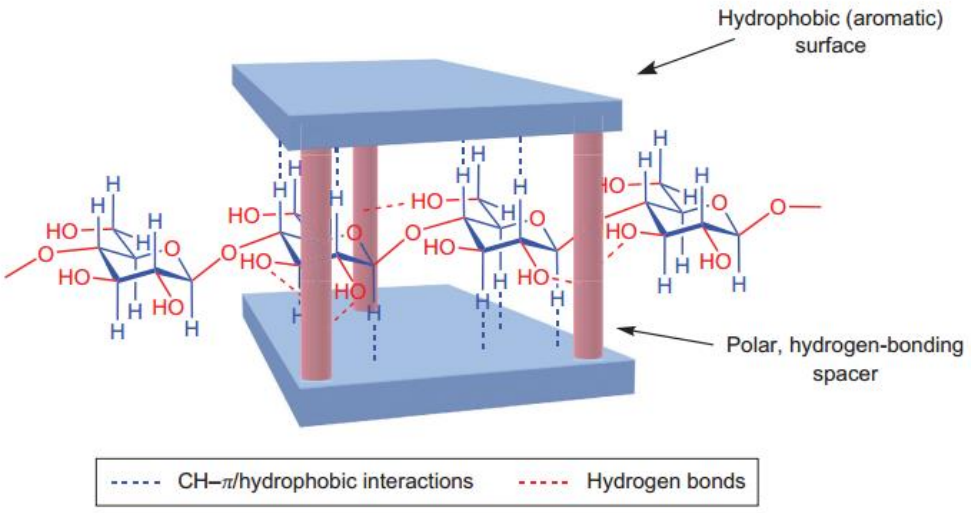


Figure 16 Three pillar pyrene based receptor **32** was shown to be have very high affinities for cellulose derived oligosaccharides (such as D-cellotetraose **33**). It is believed the sugar based polymer chain threads through the cavity until the most optimal configuration is achieved.[39]

Extensive NMR spectroscopy studies and molecular modelling suggested that the oligosaccharide chain threads through the cavity and then rests with the middle glycosidic bond in the centre of the binding site (Figure 17). The terminal monomers are then solvated by the water or have additional interactions with the dendrimer solubilising groups. Once the sugar chain threads into the receptor, interactions are possible with any of the monomers but were least favoured with the terminal sugar units according to NOESY NMR spectroscopy. Once the oligosaccharide enters the cavity it can then thread through the cavity until the most optimal interactions are achieved (in this case, the middle of the oligomer). This might explain how the resultant strong binding constants were obtained.[39]

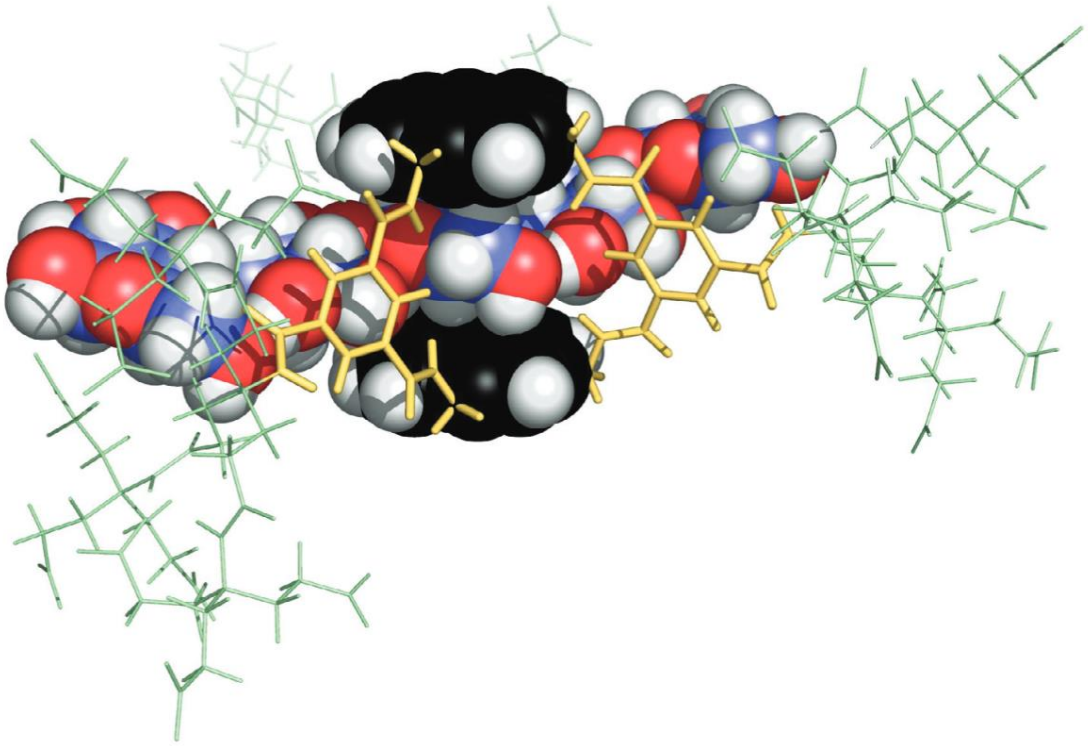


Figure 17 Molecular modelling of pyrene receptor **32** and D-cellopentaose, where optimal binding takes place at the centre of the oligosaccharide chain.[39]

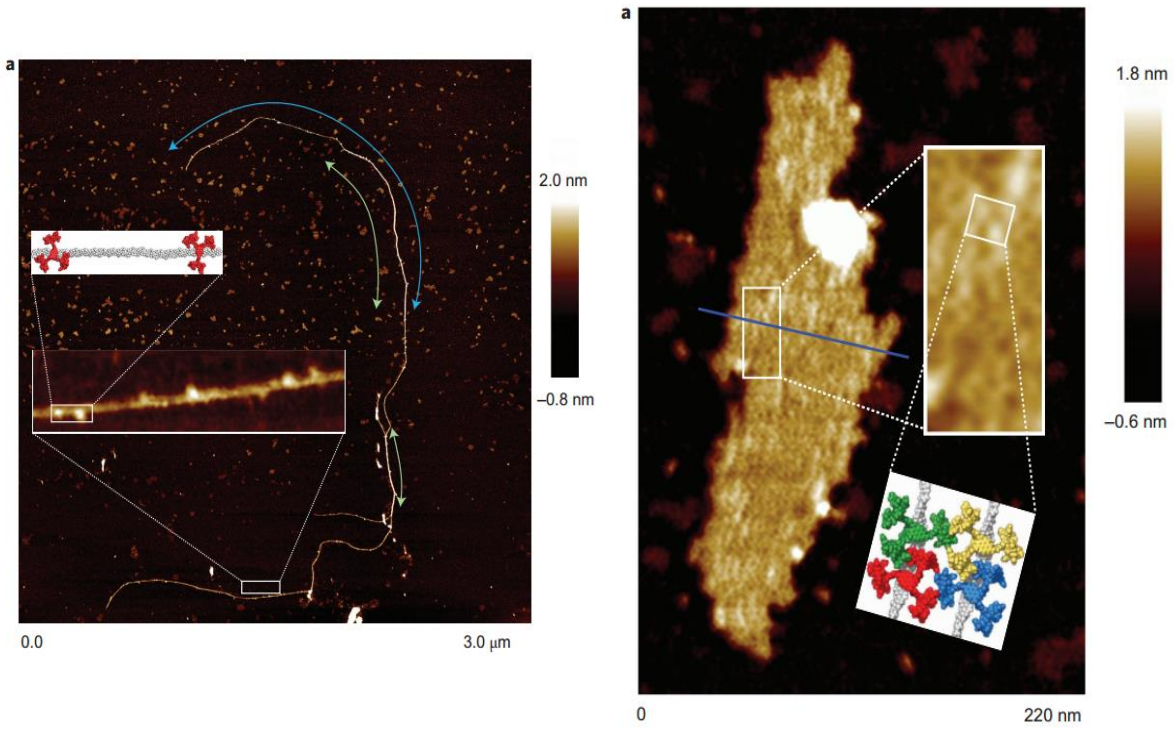
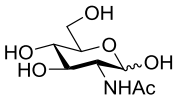
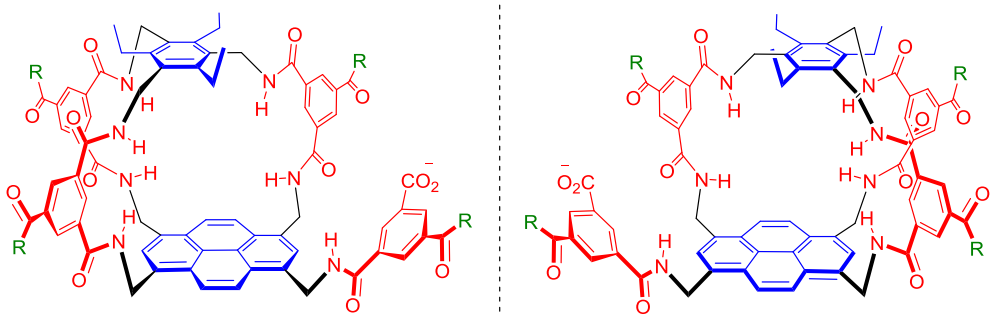


Figure 18 Analysis of receptor **32** and D-cellulose (left), and **32** and D-chitosan (right) using AFM. Observed height profiles for observed protrusions are consistent with receptor threaded onto polysaccharides.[39]

With threading of oligosaccharides now established, the threading of polysaccharides was investigated. A solution of **32** and D-cellulose in 1M NaOH, and a mildly acidic (~pH 6) solution of **32** and D-chitosan were studied by AFM for any evidence of threading (Figure 18). The height profiles for observed protrusions were consistent with the polysaccharides threading through the receptor cavity, and not with the receptor simply lying on top of the polymer surface. Control samples of the polysaccharides without receptor **32**, were shown to adopt coiled aggregates and not the strands (D-cellulose) or plates (D-chitosan) observed when receptor was also present. This would suggest that the bound receptor interferes with the self-association and folding of the substrates. D-Chitosan and receptor **32** were also found to be soluble across a broad pH range (pH 1-11), whereas the separate components are not, implying a mutual solubility was achieved as the host-guest pair. This result is the first known example of a polysaccharide-based pseudorotaxane architecture.[39]



7


R
O
N
H
O
N-H
H-N
CO2
R
O
O
O

Chapter 2 – Tetra Alkoxy Ester Anthracene Receptor

2.1 Polysaccharides as a Target Guest Substrate

D-Cellulose **8** is the most abundant organic molecule on Earth, acting as an important structural component of the primary cell wall in many green plants and algae.[2] Equally considerable is D-chitin **9** which is the second most abundant biopolymer and a major constituent of fungi and invertebrate exoskeletons.[41] This makes both polysaccharides an almost inexhaustible source of carbon when selectively degraded to their corresponding monomer units using enzymes or reagents.[42,43] D-cellulose **8** can be converted to D-glucose **4** which is valued by the fuel and food industry, whereas D-chitin **9** can be converted to *N*-acetyl-D-glucosamine **7** which has potential applications in medicine (Figure 20).[44] The main drawback to taking advantage of these materials is that cellulose and chitin do not dissolve in a neutral aqueous environment, making the polymeric material difficult to manipulate and process efficiently. Current chemical processing methods require elevated temperatures and extreme pH ranges.[45] Therefore a host molecule that assists in the solubility of these polysaccharides in a neutral aqueous environment would allow enzymes or reagents that could convert the polymeric scaffolds to their equivalent monomers under mild conditions. Receptor **32** (Figure 16) has already proven capable of threading oligosaccharides through the binding cavity but is not very synthetically accessible.[39] Receptors of alternate design that are more synthetically accessible could thread onto polysaccharide chains such as cellulose **8** and chitin **9**. It is envisaged that such receptors would form pseudo rotaxanes upon binding with the polymer chain and that several receptors per polymer could solubilise the polysaccharide under mild aqueous conditions.

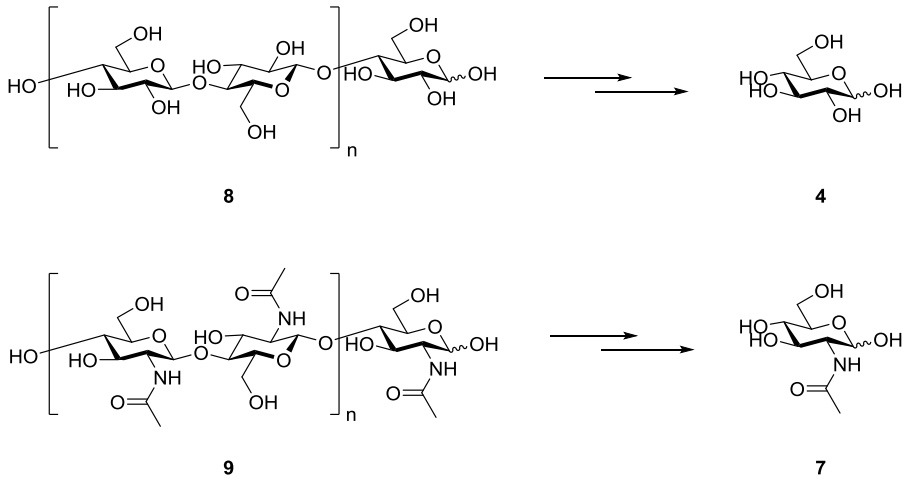


Figure 20 The most abundant biopolymers: D-cellulose **8** and D-chitin **9** and their corresponding monomers D-glucose **4** and *N*-acetyl-D-glucosamine **7**.

2.2 Receptor Design and Retrosynthetic Analysis

The main aim of this project was to design and develop novel receptors that are capable of binding oligo and polysaccharides by threading of the polymer chain through the cavity. As shown before, receptor **32** showed strong binding affinities for oligosaccharides derived from β-1,4 polysaccharides such as cellulose and chitosan, where the sugar chain was believed to thread through the central cavity (Figure 21).[39] This new found activity was not observed for previous systems, such as for receptors **24** and **25**.[33,35] This was postulated to be due to **32** having a slightly larger cavity which allows the terminal sugar to enter and then thread through until optimal interactions are achieved at the central glycoside bridge. Despite the promising results obtained for **32**, the receptor itself was not trivial to synthesise as it requires a number of synthetic steps and functional group transformations (>20 steps), not all of which proceed smoothly.

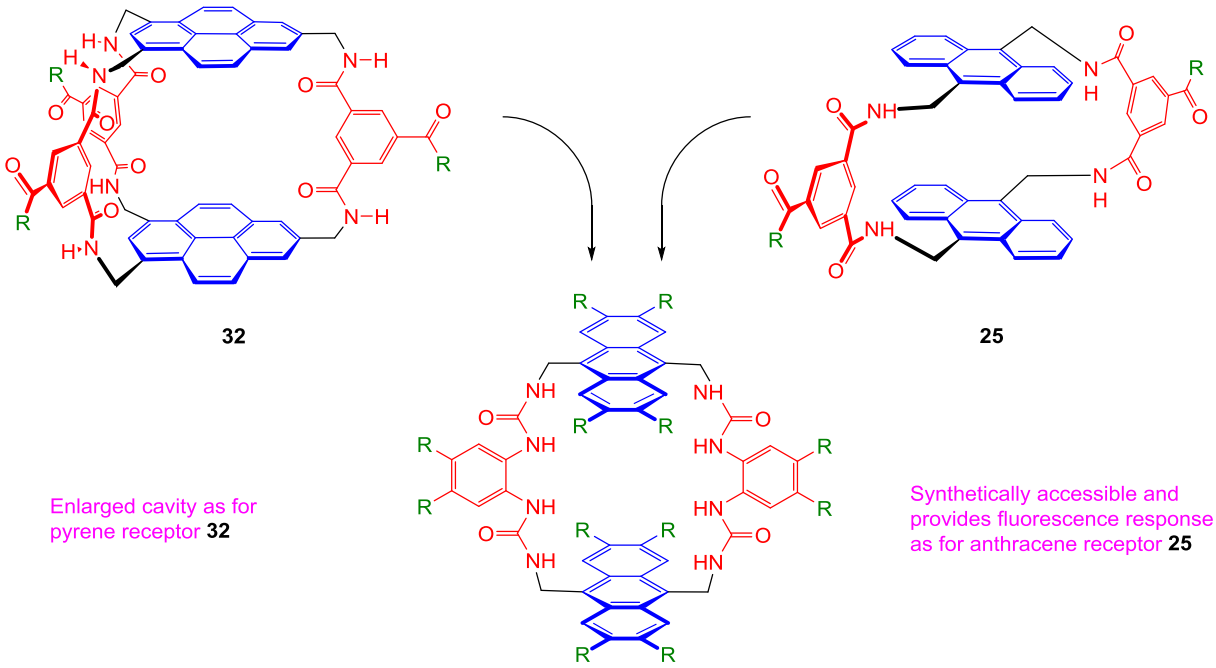


Figure 21 Simplification of **32** to tetraurea **35** which takes inspiration from **25**, where R indicates modification points that could aid solubility or binding in water. Receptor **25** shows very weak or no binding to oligosaccharides in water, believed to be due to a small cavity size. Receptor **32** is not easy to synthesise.

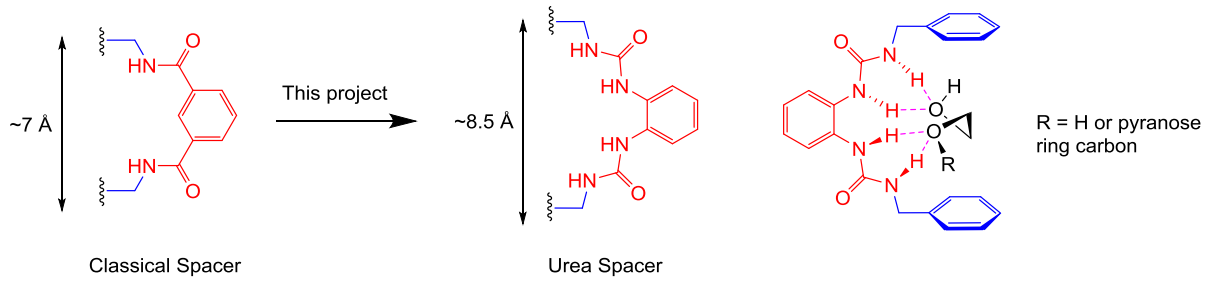


Figure 22 Comparison of bis-amide (left, 10 bonds) and bis-urea (centre, 11 bonds) spacer units. Hypothesised hydrogen bonding array provided by the bis-urea spacer towards a diol motif commonly found in carbohydrates (right).

Taking inspiration from the relatively simple synthetic route towards **25** (when compared to **32**), the design for tetraurea **35** was proposed (Figure 21). The increased cavity size of **35** was to be achieved through lengthening of the polar spacer units by one bond length, progressing from a bis-amide to a bis-urea motif (Figure 22). While the bis-urea spacer enlarges the cavity, it would also effectively double the number of H-bond donors that are directed towards the hydroxyl groups of the sugar substrate within the binding site (Figure 22). The favourable intermolecular hydrogen bonds between the ureas and the oxygen lone pairs of the sugar were observed through Monte Carlo Molecular Mechanics (MCMM) which modelled the complex of the receptor and cellotetraose (Figure 23), which demonstrated hydrogen bonding with the central hydroxyls of the oligosaccharide chain once in the cavity. Variation of the H-bond donors has not yet been investigated for these temple style receptors and so the impact of this change upon receptor conformation and binding to smaller substrates (such as D-glucose) will also be explored. The bis-urea spacer unit also holds the aromatic surfaces further apart: approximately 8.5 Å, which is close to ideal for accommodating a carbohydrate. This increased distance is also significantly larger than that required for π-π stacking of aromatics (~7 Å) and was predicted to disfavour the binding of aromatic guests. The strong binding of small aromatics is a known problem for the selectivity of carbohydrate receptor **25**, which has a smaller distance between the aromatic surfaces. The anthracene units of **35** run parallel with the polysaccharide to maximise CH-π interactions, but also provide a secondary medium to observe binding through fluorescence/UV-vis spectroscopy.

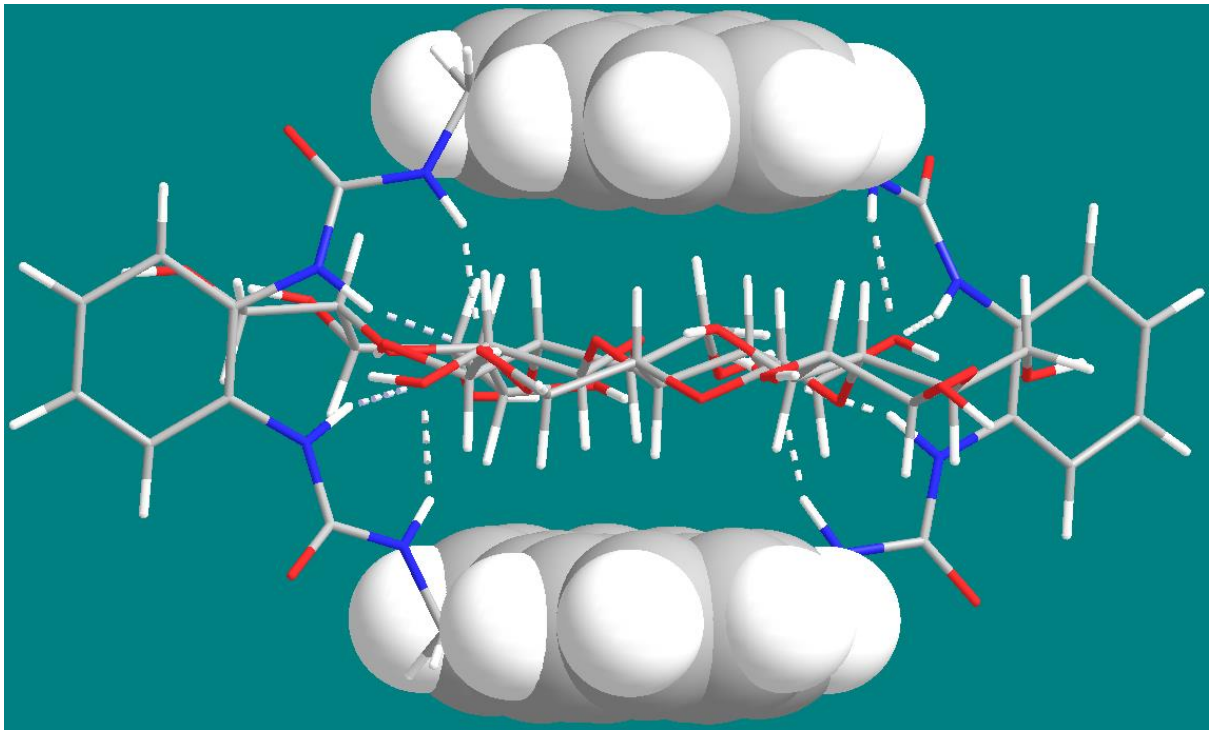
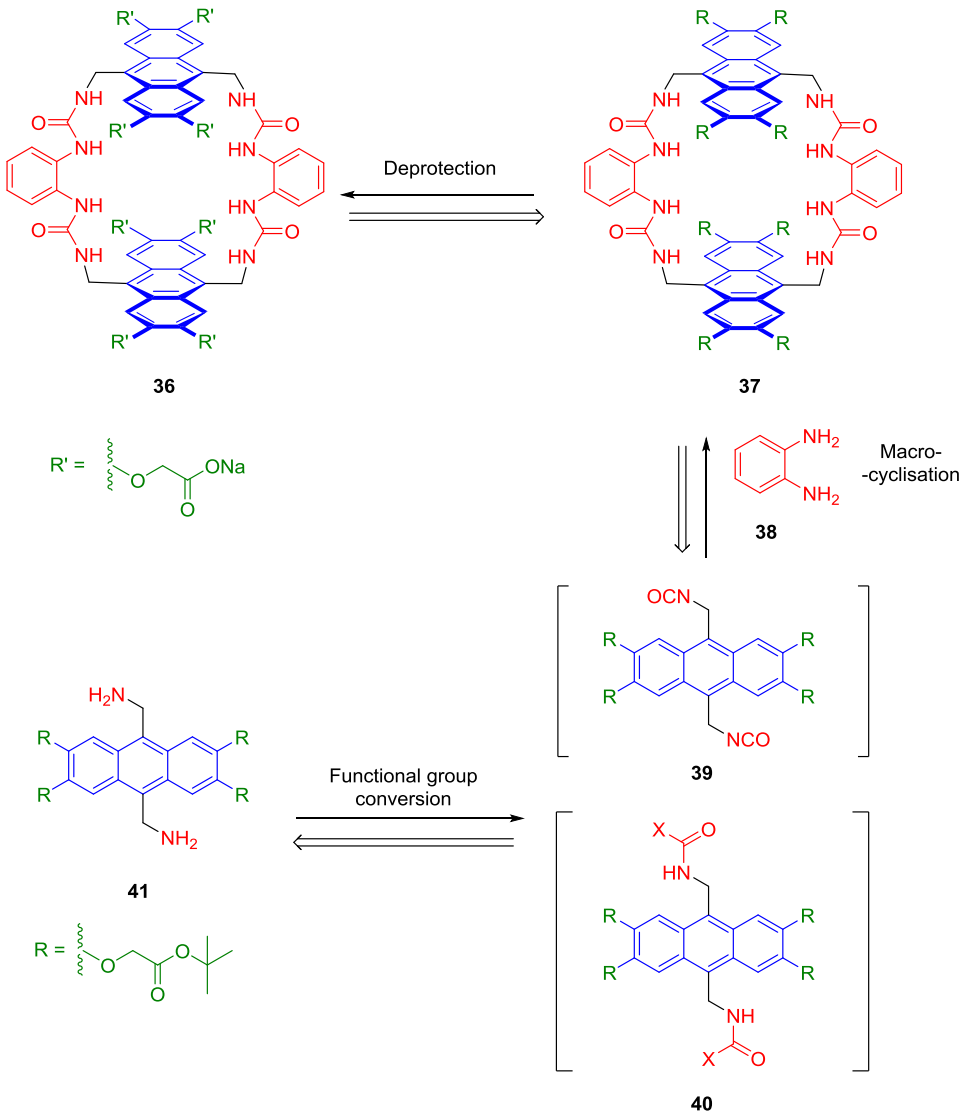


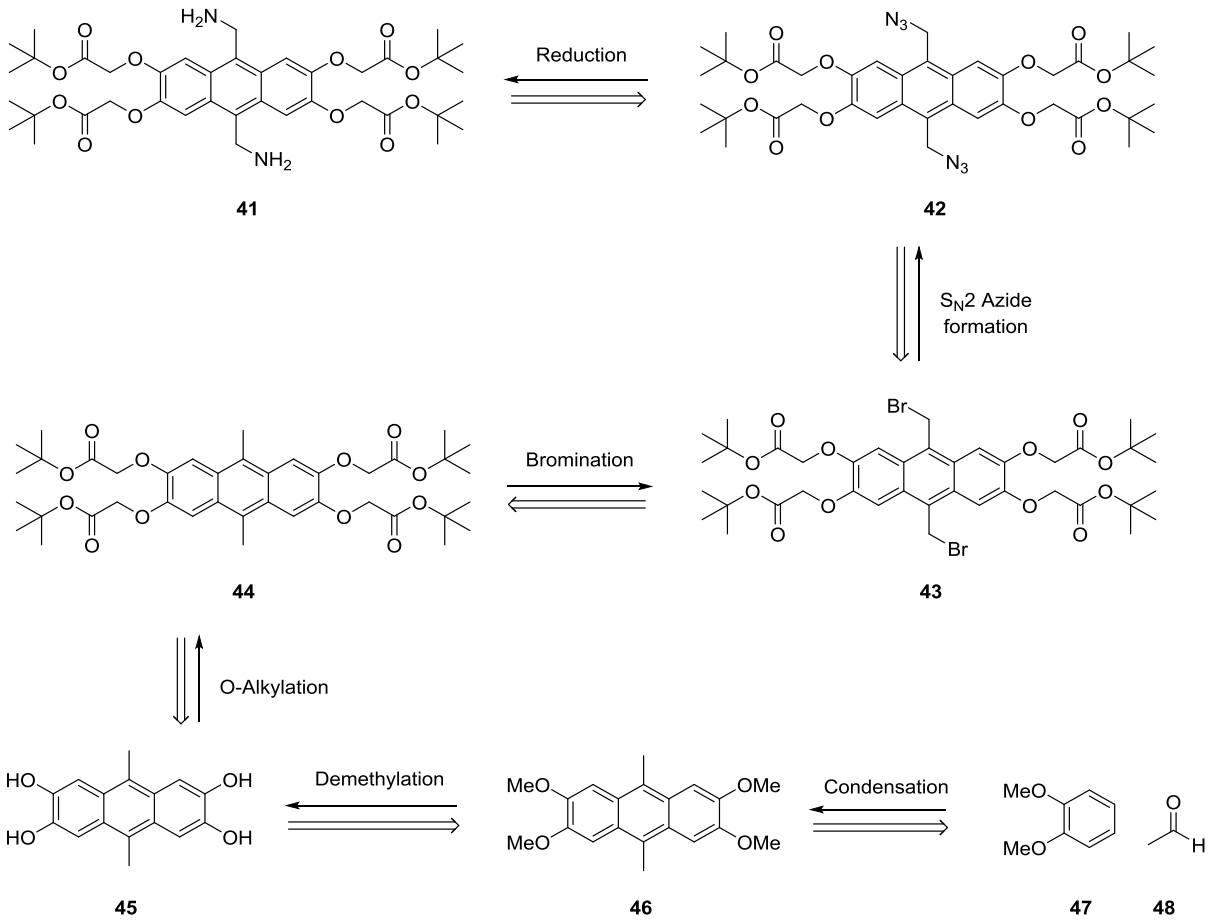
Figure 23 Monte Carlo Molecular Mechanics (MCMM) of the tetraurea anthracene **35** and D-cellotetraose **33**, showing the urea hydrogen bonds and the central hydroxyls near the glycosidic bond (dashed lines). Force field used: OPLS2005 with aqueous GB/SA solvation.

Initial synthetic efforts targeted receptors of the design **36** (Scheme 1). These feature four alkoxy carboxylate groups per anthracene, eight in total per receptor, to aid solubility in aqueous media. The classical temple systems, such as **25**, featured the solubilising groups on the polar linkers. These anthracene systems were water soluble with as few as six carboxylate groups, so it was rationalised that eight carboxylates for **36** would provide sufficient solubility of the receptor in water. Attachment of the solubilising groups to the anthracene (in the case of **36**) was to allow development of methodology that would give access to the bis-urea spacer unit.



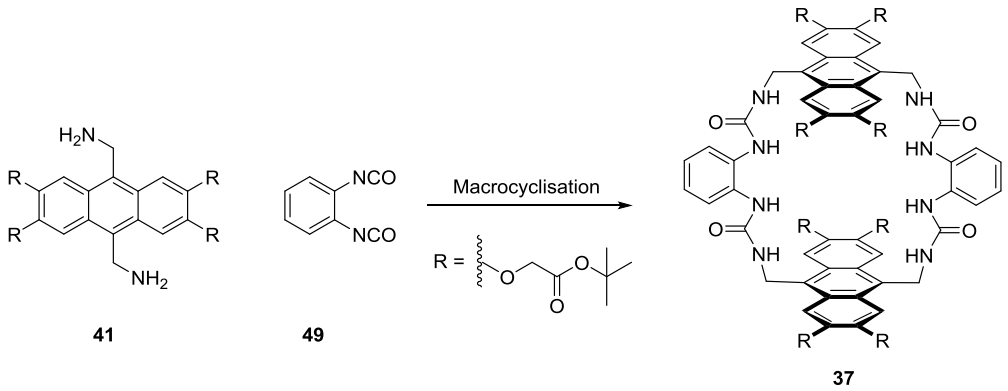
Scheme 1. Retrosynthetic analysis of tetraurea macrocycle to anthracene diamine (X = suitable leaving group, e.g. OC_6F_5, Cl etc.).

The key step in the synthesis of **36** was formation of the *tert*-butyl protected macrocycle **37**, which involves cyclisation of 1,2-phenylene diamine **38** and an activated amide/carbamate **40** or isocyanate intermediate **39** to install the urea functionality (Scheme 1). These isocyanates or activated amides/carbamates are synthesised after functional group conversion from the corresponding anthracene diamine **41**. A synthetic route towards **41** had already been established previously in the Davis group (Scheme 2), although yields for reduction of azide **42** to amine **41** were poor (~30%).[46] Therefore further optimisation of this step was considered desirable during the synthesis to maintain efficiency of material. Once the protected macrocycle **37** can be synthesised, deprotection of the *tert*-butyl esters to the carboxylate *via* acid and subsequent neutralisation to neutral pH yields the water soluble macrocycle **36**.



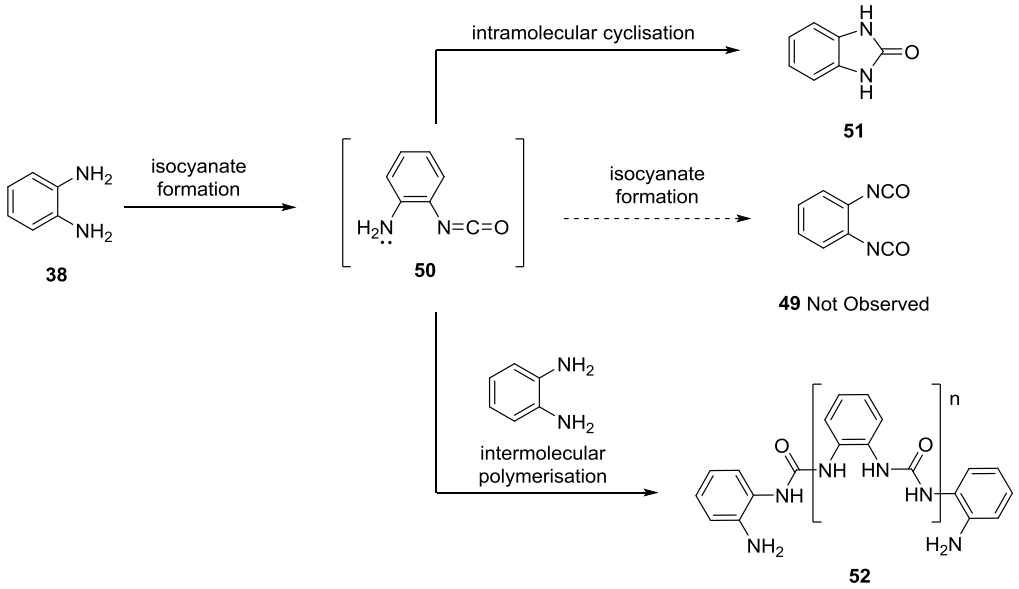
Scheme 2. Retrosynthetic analysis of the tetra ester anthracene diamine **41**

It should be noted that the reverse route which features a 1,2-diisocyanatobenzene **49** and the anthracene diamine **41** (Scheme 3) was hypothesised to not be viable, due to the inherent instability of **49**. Literature regarding **49** was scarce but suggests issues with polymerisation and intramolecular cyclisation (Scheme 4).[47] Unpublished work within the Davis group reinforces this hypothesis, as Boc protected carbamate **53** was shown to unintentionally cyclise through reaction with the adjacent imidazole urea (Scheme 5).[48] Both of these functional groups are less reactive than the equivalent amine and isocyanate present on intermediate **50**. It was therefore predicted that similar, if not worse issues would be faced trying to synthesise **49**.



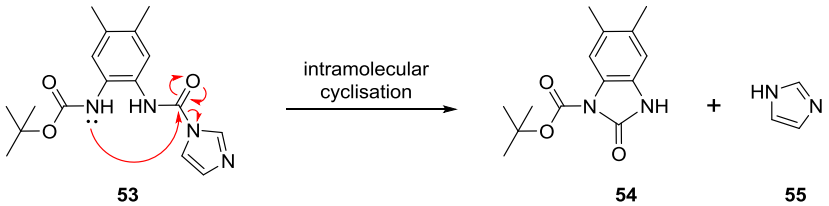
41 **49** **37**

Scheme 3. Macrocyclisation with anthracene diamine **41** and 1,2-diisocyanatobenzene **49**.



38 **50** **51** **49** Not Observed **52**

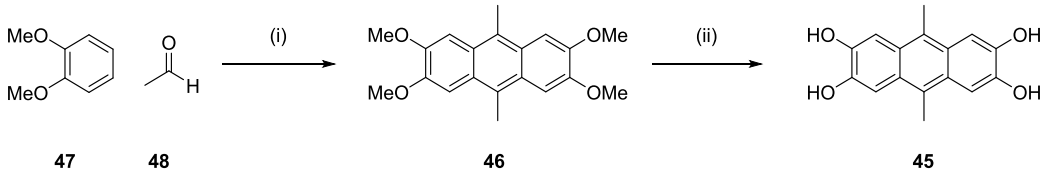
Scheme 4. Reported attempts to make 1,2-diisocyanatobenzene **49** only observe cyclic **51** and polymeric ureas **52**.[47]



53 **54** **55**

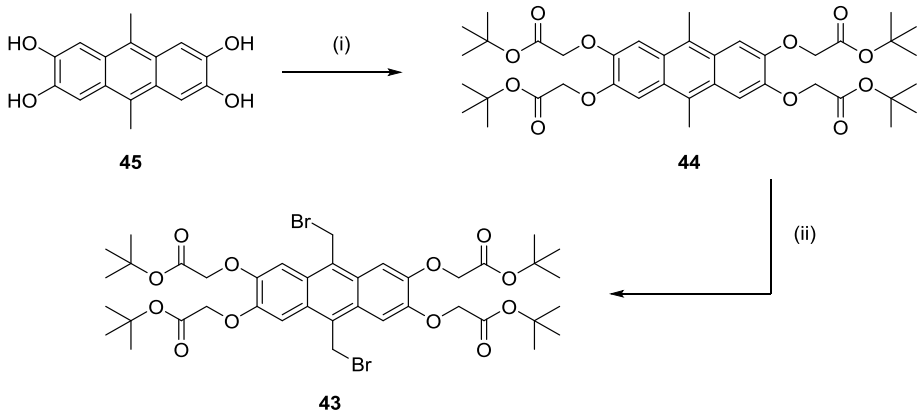
Scheme 5. Cyclisation of *tert*-butyl carbamate **53** onto imidazole urea to give cyclic urea **54** and imidazole **55**, as previously observed by the Davis group.[48]

2.3 Synthesis



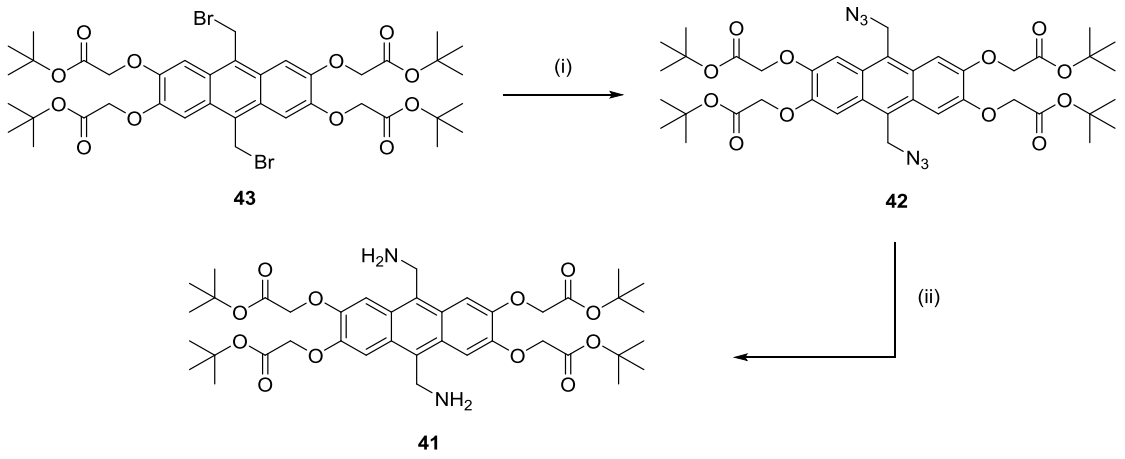
Scheme 6. (i) H_2SO_4, MeCN, -10 °C, 2h, 61%; (ii) BBr_3, CH_2Cl_2, reflux, 16h, 79%.

Initial construction of the anthracene scaffold **46** was carried out by a condensation between 1,2-dimethoxybenzene **47** and acetaldehyde **48** in concentrated sulfuric acid (Scheme 6). It was found that maintaining a temperature below 0 °C was crucial to preventing polymerisation and thus increase the yield. The resultant crude mixture obtained was highly insoluble and proved difficult to purify. Purification by flash column chromatography using chloroform as an eluent was somewhat successful, although had to be performed in batches on large scale due to insolubility. Demethylation of **46** to tetra alcohol **45** was then achieved in good yield by addition of boron tribromide and then refluxing in dichloromethane.



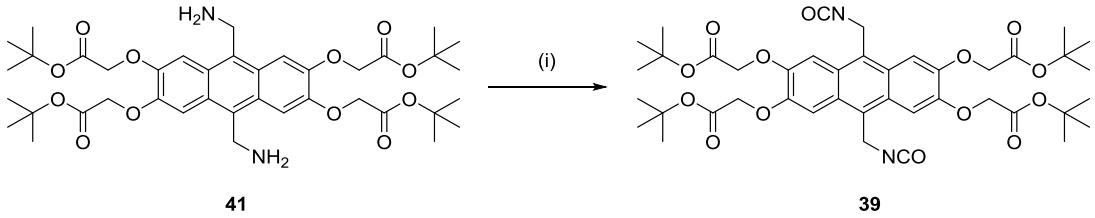
Scheme 7. (i) *tert*-butyl bromoacetate, K_2CO_3, THF, reflux, 16h, 67%; (ii) NBS, ABCN, DCM, reflux, 1h, 98%.

Alkylation of **45** with *tert*-butyl bromoacetate in basic conditions then delivered tetra ester **44** in good yield after column chromatography (Scheme 7). A radical bromination of **44** with N-bromosuccinimide and radical initiator ABCN then afforded the dibromide **43** in quantitative yield. No further purification was necessary after extraction of the bromide using dichloromethane and washing with aqueous base to remove residual succinimide.



Scheme 8. (i) NaN_3, MeCN, reflux, 3h, 98%; (ii) PMe_3, H_2O, THF, rt, 16h, 98%.

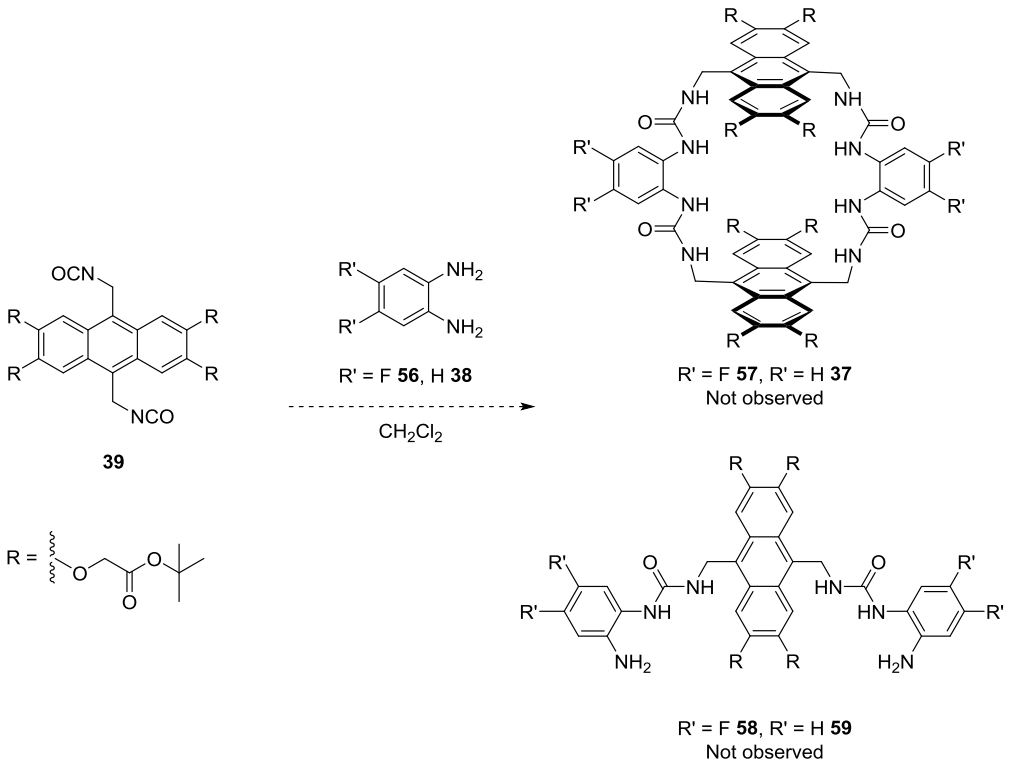
The corresponding diazido anthracene **42** was then obtained through S_N2 substitution of the bromide **43** with sodium azide, which was obtained in a near quantitative yield (Scheme 8). The crude azide **42** was then converted to the desired diamine **41** in a quantitative yield *via* a Staudinger reaction using trimethylphosphine in THF. It was found that using a great excess (up to 20 equivalents) of phosphine was needed to drive the reaction to completion and overcome the poor yield previously reported for this transformation. The residual trimethylphosphine oxide was also easily removed by sublimation *via* freeze-drying, meaning no column chromatography was necessary.



Scheme 9. (i) Triphosgene, $NaHCO_3$ (aqueous, 0.1 M), CH_2Cl_2, 0 °C, 30 min, 96%.

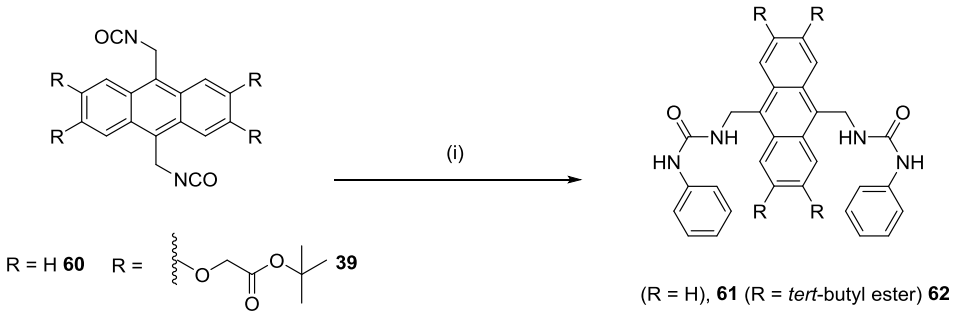
With diamine **41** in hand, the next stage was to convert it to either an isocyanate **39** or activated amide/carbamate **40**, which would allow access to the macrocycle upon reaction with 1,2-phenylene diamine **38** (Scheme 1). It was decided to pursue the isocyanate route initially, as the group had previously experienced success in their use for urea bond formation in the synthesis of anion transporters. Anthracene diamine **41** was then exposed to previously reported conditions that make use of a biphasic solvent system (Scheme 9), which afforded diisocyanate **39** in an excellent yield with no further purification necessary. A very clear shift downfield of the methylene peak adjacent to nitrogen was observed in the ^1H NMR, but the isocyanate carbon itself proved too weak to be observed by ^{13}C NMR. The NCO functional group showed a reasonably strong characteristic infra-red stretch at

~2250 cm^{-1} however, and this was used to confirm the presence of the isocyanate functional group. While the product obtained was a solid, it showed slow degradation in air and hydrolysis in solution over time. Storage under nitrogen in the solid state at 0 °C was necessary if kept for periods of time longer than a week, but generally the diisocyanate **39** was used instantly for the next step after isolation.



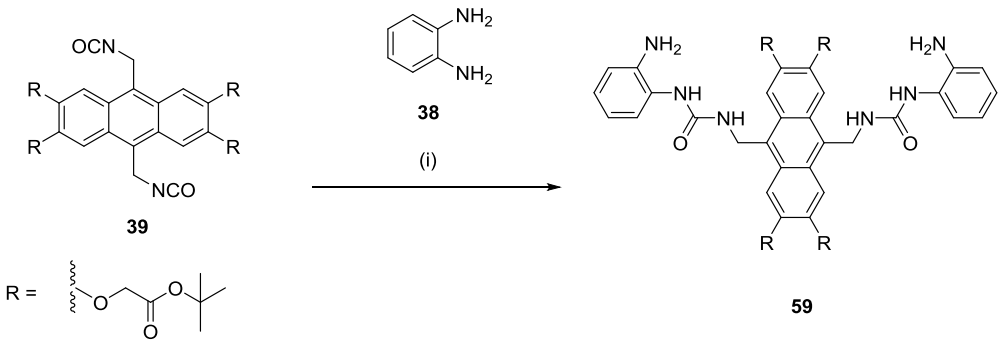
Scheme 10. Attempted Macrocyclisation with diisocyanate **39** and two different diamino benzenes **38** and **56**.

Initial attempts to cyclise **39** with 1,2-diamino-4,5-difluorobenzene **56** were unsuccessful (Scheme 10, R' = F), with none of macrocycle **57** observed in the crude reaction mixture (by ^{1}H NMR or mass spectrometry) or after purification by reverse phase HPLC. The fluorinated diamine **56** was initially chosen with the intent that the fluorine would allow easy identification and characterisation of the product by ^{1}H and ^{19}F NMR. It was afterwards postulated that the negatively inductive effects of fluorine could reduce the nucleophilicity of the diamine, especially once one of the amines had formed a urea to make the half macrocycle. The reaction was then repeated with 1,2-phenylene diamine **38** (R' = H, Scheme 10), but again no macrocycle **37** was observed or isolated. Surprisingly, efforts to synthesise the half receptor **59** also failed using one equivalent of the diisocyanate **39** and 2 equivalents of diamine **38** (Scheme 10). This lack of reactivity towards the isocyanate was unexpected. It was initially rationalised that either the *tert*-butyl esters make the approach towards the isocyanate too hindered for the diamine or that once a monourea is formed the resultant molecule becomes less soluble in the reaction medium and is very slow to progress to the full macrocycle.



Scheme 11. (i) Aniline, CH_2Cl_2, rt, 16h, **61** = 85%, **62** = 82%.

To probe this apparent unreactive nature of the isocyanate, test reactions were undertaken to simulate formation of the half macrocycle. Simple diisocyanate anthracene **60** (R = H, Scheme 11) was synthesised from the corresponding diamine, which was prepared from literature procedure.[49] Conversion of this anthracene diamine to the isocyanate **60** was carried using the same procedure as for **39** (Scheme 9), and was then reacted with aniline in dichloromethane. After stirring for 16 hours bis-urea **61** precipitated out in an excellent yield. **61** proved to be highly insoluble in most organic solvents, except in more polar media such as DMSO or pyridine and this would indicate that insolubility of the macrocyclic intermediates could be a problem. The formation of this bis-urea however does suggest that the isocyanate was sufficiently electrophilic to react with the diamine, at least to form the half macrocycle, and that perhaps the steric hindrance of the *tert*-butyl esters could also be an issue. To probe this hypothesis, **39** was then reacted with aniline in dichloromethane. Resultant isolation of **62** then appeared to disprove this, but its high degree of insolubility like **61** would agree with the suggestion that poor solubility of urea intermediates in the macrocyclisation could be problematic.



Scheme 12. (i) 1,2-phenylenediamine **38**, CH_2Cl_2, rt, 1h, 87%.

Synthesis of the half macrocycle **59** was now a priority, so that its solubility could be better investigated and thus a more appropriate solvent be chosen for the macrocycle formation. Addition of diisocyanate **39** to a large excess of diamine **38** (~20 equivalents, with previous attempts using <3 equivalents) afforded half receptor **59** in an excellent yield (Scheme 12). The large excess of diamine **38** was designed to maximise the exposure of isocyanate **39** to diamine and limit any potential reaction of half

receptor with isocyanate. As expected, **59** also proved very insoluble in most organic solvents, but demonstrated good solubility in more polar solvents such as DMSO, DMF and pyridine.

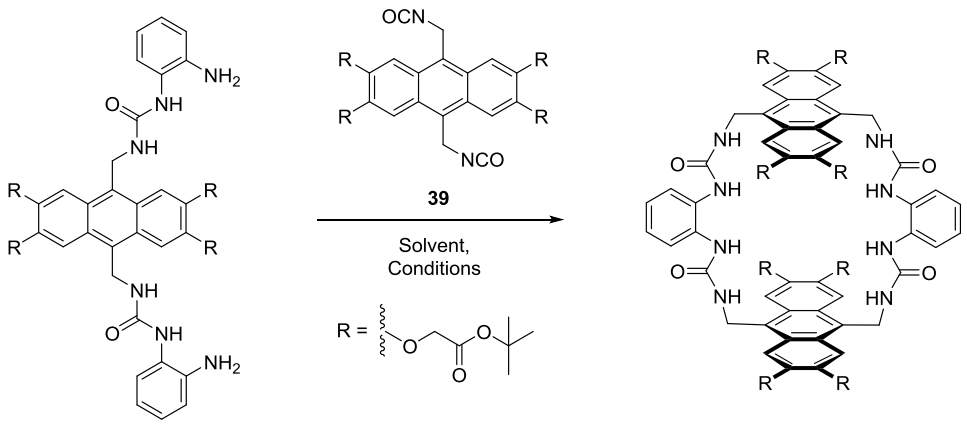


Table 1. Solvents and conditions screened for macrocyclisation of **59** and **39**.

Entry	Solvent	Isolated Yield (%)	Conditions	Comments
1	CH_2Cl_2	0	rt, 3 days	Minor isocyanate hydrolysis, half receptor recovered.
2	CH_2Cl_2	0	DIPEA, rt, 1 day	Same as entry 1, isocyanate hydrolysed faster.
3	CH_2Cl_2	0	reflux, 7 days	Same as entry 1
4	THF	0	reflux, 3 days	Same as entry 1, half receptor less soluble.
5	DMF	-	rt, 16h	Consumption of half receptor, no separation by HPLC.
6	CH_2Cl_2:pyridine (5:1)	<10%	rt, <3h	Macrocycle observed in mass spectrum of crude reaction mixture.

Several solvents and conditions were screened to promote formation of the target macrocycle **37** from half receptor **59** and diisocyanate **39** (Table 1). All reactions were performed at relatively high concentrations (5 mM) to promote any sign of reactivity, whereas macrocyclisation formations typically use more dilute conditions, such as 1-0.1 mM, to favour intramolecular cyclisation. Initially dichloromethane was chosen as a solvent to confirm the hypothesis that solubility of **59** was the limiting factor (entry 1, Table 1). Indeed, only a small amount of hydrolysis of isocyanate **39** was observed, with the half macrocycle **59** being recovered afterwards. Addition of the base DIPEA (entry 2, Table 1) appeared to have no effect, except to accelerate the hydrolysis of the isocyanate **39**. Refluxing the reaction mixture in dichloromethane (entry 3, Table 1) still showed the presence of **39** and **59** after one week. Interestingly, addition of hexyl isocyanate to half receptor **59** in dichloromethane gave the corresponding tetraurea demonstrating that the half receptor was sufficiently nucleophilic. This might suggest any macrocyclic intermediates formed during entries 1-3 are probably too insoluble to allow the reaction to proceed to completion. More polar solvents were then screened, but **59** appeared even less soluble in THF than dichloromethane and even upon refluxing no consumption of **59** was seen. Half macrocycle **59** proved to be highly soluble in DMF (entry 5, Table 1) and was fully consumed after 16 hours at room temperature. Subsequent purification by preparative reverse phase HPLC proved non-trivial compared to previous macrocyclic systems in the group and it was eventually evident that a complex mixture of products had been obtained.

Half receptor **59** was already known to be very soluble in pyridine which was to be the next solvent system to be investigated, and it was eventually found that addition of half receptor **59** in pyridine to a dilute solution of isocyanate **39** in dichloromethane (final pyridine to CH_2Cl_2, 1:5) showed complete consumption of both starting materials in under 3 hours, with new fluorescent compounds clearly visible by TLC. While this improved rate of reactivity could be attributed to improved solubility, pyridine is also known to act as a nucleophilic catalyst for acylation and could be activating the isocyanate *via* a pyridinium intermediate – thus increasing the electrophilicity at the isocyanate carbon, although some previous work disputes this hypothesis and suggests pyridine is merely acting as a base.[50] Analysis of the crude reaction mixture by mass spectrometry confirmed the presence of the target macrocycle **37**. Initial purification by column chromatography and then analysis of this mixture by reverse phase HPLC suggested there were 3 major products (Figure 20) with similar retention times: potentially all macrocyclic species of varying ring sizes.

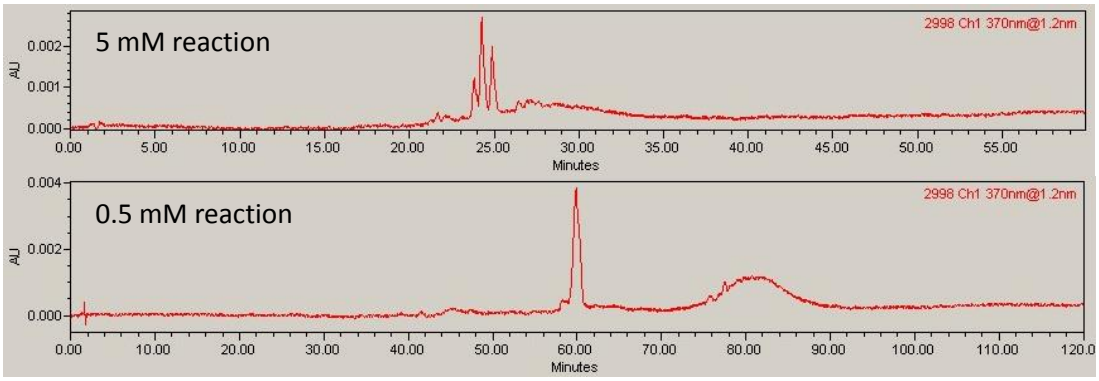
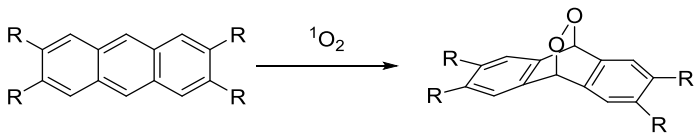


Figure 24. Reverse phase HPLC traces of the partially purified reaction mixtures to form macrocycle **37**, using reaction conditions outlined in entry 6, Table 1. The difference in reaction concentration can be seen, with a much cleaner reaction being obtained upon dilution of the reaction from 5mM (top) to 0.5 mM (bottom). Detection wavelengths of 370 nm were used to detect anthracene. Each run was performed using different eluent gradients, hence the differing retention times. Stationary phase: analytical scale C18, flow rate: 1 mL/min, eluents: acetone/water.

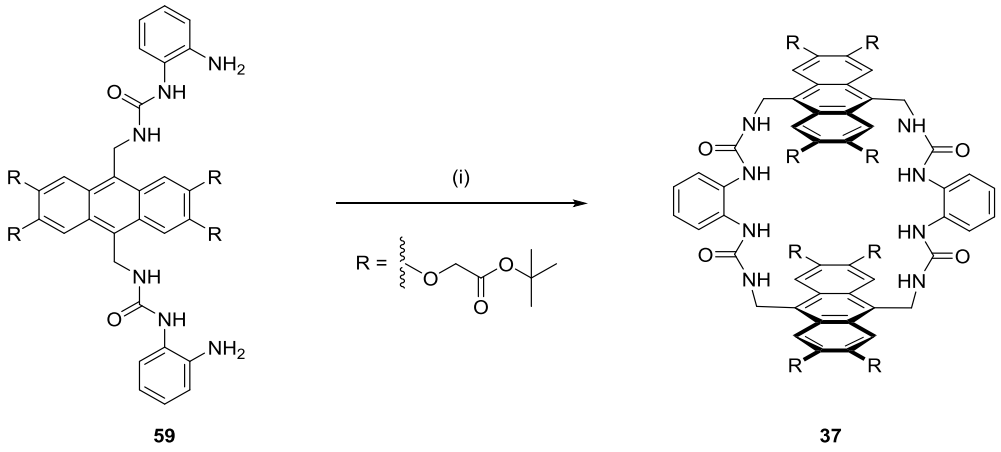
These test macrocyclisations were performed relatively concentrated (5 mM), so dilution of the reaction mixture (0.1-1 mM) was expected to favour the intramolecular cyclisation that leads to the target macrocycle **37**. Indeed, such a result was obtained upon dilution to 0.5 mM, as seen in the much cleaner HPLC trace obtained (Figure 24). Macrocycle **37** was then isolated by preparative HPLC in a moderate yield of 40%. Tetraurea **37** proved more soluble than half receptor **59** but displayed aggregation in most deuterated solvents during characterisation, even upon dilution. Using $(CD_3)_2CO$ as a solvent was found to give the ^1H NMR spectrum with the sharpest signals, although some peaks relating to the spacer aromatics were still heavily broadened. This could be due to slow conformational exchange caused by hydrogen bonding taking place between the ureas on the spacer units.



Scheme 13. Diels-Alder addition of singlet oxygen to anthracene.

Also evident was partial oxidation of one or both of the anthracene units over time, a result of a Diels-Alder addition of singlet oxygen that forms an irreversible peroxide bridge across the central aromatic ring (Scheme 13).[51] The mono-oxidised and bis-oxidised species were both observed using mass spectrometry, which would suggest they are relatively stable being able to survive the high energy environment during mass spectrometry. Such oxidation was observed to get progressively worse in a matter of hours during characterisation. Previous studies in the literature have demonstrated that the lifetime of singlet oxygen decreases significantly in protic solvents (such as methanol and water) due

to H-bond stabilisation, and that deuteration increases the lifetime.[52,53] Therefore aprotic NMR solvents such as $CDCl_3$ and $(CD_3)_2CO$ are predicted to show increased rates of oxidation of anthracenes. Oxidation of anthracene has been prevalent in previous electron rich anthracene based systems in the Davis group[46], but was not observed in this system until now. It was also not clear when this oxidation had taken place, so future efforts in synthesising **37** made sure to remove all oxygen from the system and minimise its time spent in solution. This resulted in removal of the flash column from the purification step and purification of the crude product only by preparative reverse phase HPLC as soon as the reaction had reached completion.

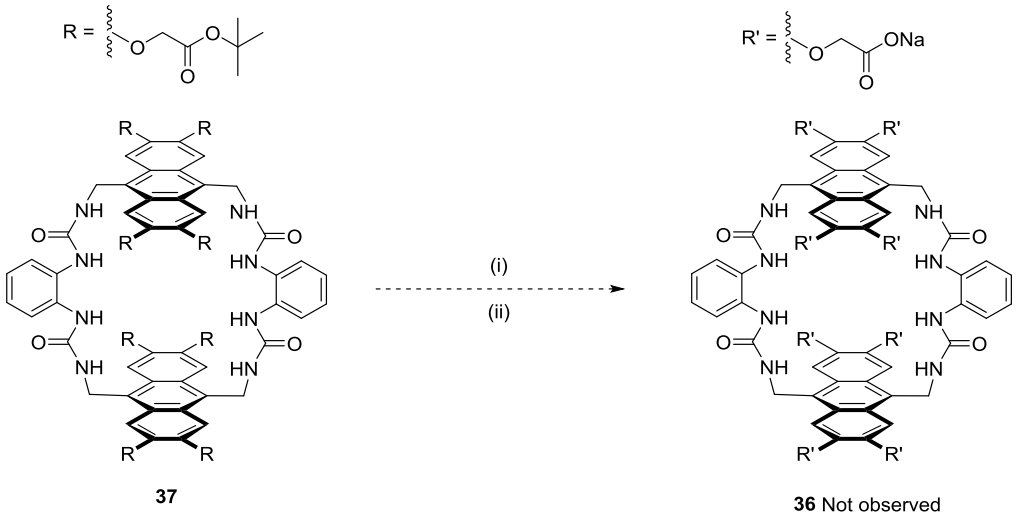


Scheme 14. (i) 43, CH_2Cl_2 and pyridine (10:1), 16h, rt, 56%.

Further optimisation led to a reduction of pyridine used, from 5:1 (CH_2Cl_2:pyridine) to 10:1. This final optimised procedure (Scheme 14) afforded **37** in 56% yield, a relatively high yield when compared to previous macrocyclisations in the group (typically <20%).[35,36] This might be due to the much faster cyclisation rate for **37** than for previous systems, thus preferentially forming the desired macrocycle before any side reactions or degradation of starting materials can occur. It was also be noted that a macrocyclisation (i.e. 2+2) between diisocyanate **39** and 1,2-phenylene diamine **38** in a 10:1 mixture of dichloromethane and pyridine was attempted. This yielded mostly half receptor **59** (>70%) and very little macrocycle **37** (<10%). It was not immediately clear why more of macrocycle **37** was not formed, but it was decided that the synthetic route *via* the half receptor was the best solution regardless, and the most efficient practically and chemically.

2.4 Attempts to access the water soluble macrocycle and investigations into macrocycle stability

Now that the protected macrocycle **37** was in hand, deprotection of the *tert*-butyl esters affords the water soluble macrocycle **36** (Scheme 15). Once in water, oxidation of anthracene was expected to be no longer be an issue due to the stabilisation of singlet oxygen through H-bond donation from water.[52,53] However, initial deprotection attempts using TFA (50% v/v) in dichloromethane, followed by neutralisation to pH 7 with aqueous sodium hydroxide, gave a fluorescent water soluble product that did not correspond to **36.**



Scheme 15 (i) TFA (50% v/v), CH_2Cl_2, rt, 16h; (ii) NaOH (aq), H_2O. Desired macrocycle **36** not observed.

The ^1H NMR spectrum in D_2O of the major products was initially hard to interpret, but it was clear the product was no longer symmetrical. Particularly evident were an increased number of signals in the aromatic region (δ 7-7.8 ppm), as well as complex new signals in the aliphatic regions (δ 0.5-4 ppm). The broad residual water peak at 4.79 ppm obscured the region where the anthracenyl and carboxyl methylene signals are usually observed, which meant that all structural information was to be garnered from the aromatic region of the spectrum. Interestingly, sequential additions of CD_3OD resulted in an increase in resolution of signals in the aromatic region, until at 100% CD_3OD the spectrum in Figure 25 was obtained. The majority of NMR analysis was thus carried out in CD_3OD due to this greater clarity.

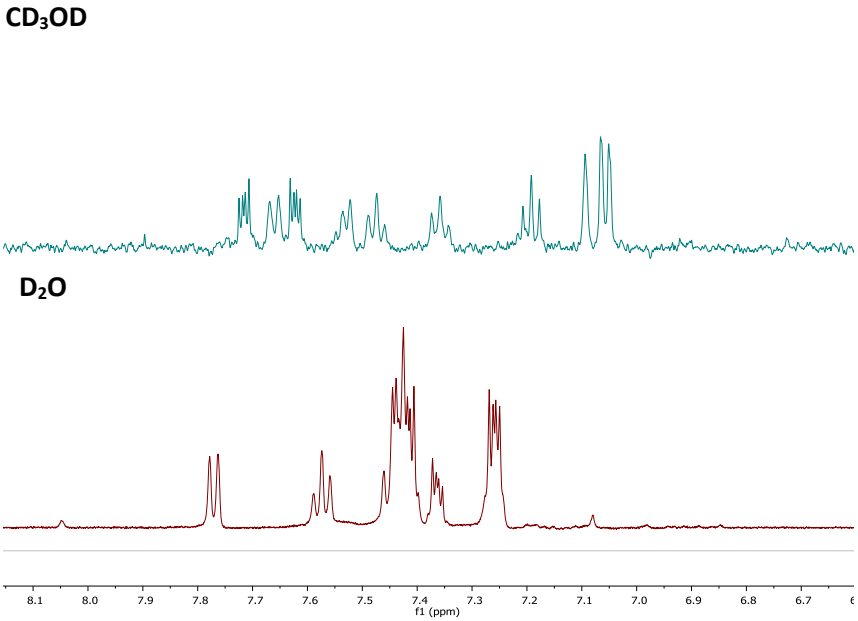


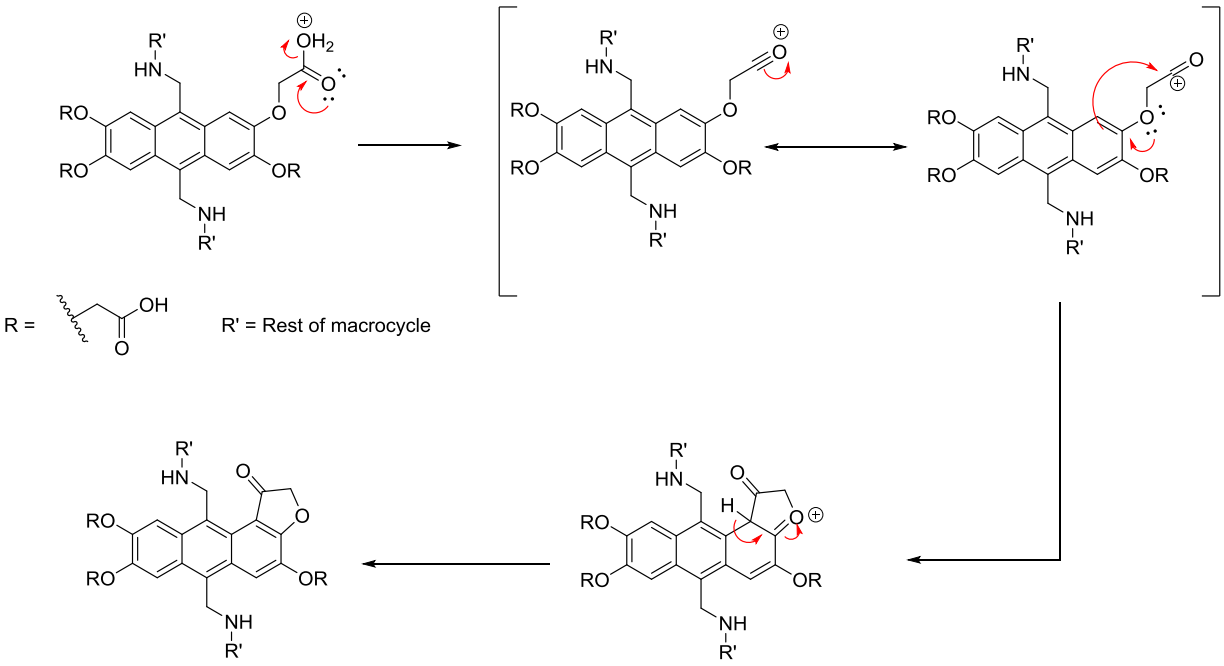
Figure 25 Partial ^{1}H NMR spectra (400 MHz, 298 K) of the major product using conditions in Scheme **15**, focusing on the aromatic region (δ 6.7-8.1 ppm). Increased resolution of peaks can be seen using CD_3OD (top) as a solvent when compared to using D_2O (bottom) as a solvent.

Extensive NMR analysis using 1D and 2D NMR characterisation techniques confirmed that the deprotection had reached 'completion', with no evidence for any *tert*-butyl groups to still be present. The increased resolution of the spectra in methanol compared to water could infer intermolecular aggregation or slow conformational exchange was occurring in water. This was disproven however when dilution (to less than 0.1 mM) and variable temperature (20-80 °C) studies in both solvents suggested the asymmetry was not a result of aggregation or of conformational origin. Another hypothesis was that the *tert*-butyl cations generated during the deprotection were potentially interfering with the reaction.[54] Commonly used cation scavenger triethylsilane was tested but the same complex product was consistently obtained. Any reaction products were not detected by mass spectrometry due to the presumed difficulty of accelerating ions derived from polycarboxylic acids, even at acidic pH in a variety of solvent mixtures. Derivatisation of the presumed carboxylates to methyl esters was intended to enable purification or ease characterisation by NMR and mass spectrometry. However, conversion to the methyl ester *via* Fischer esterification was unsuccessful and only afforded a more complex NMR spectrum, possibly due to incomplete reaction or eventual oxidation of anthracene.

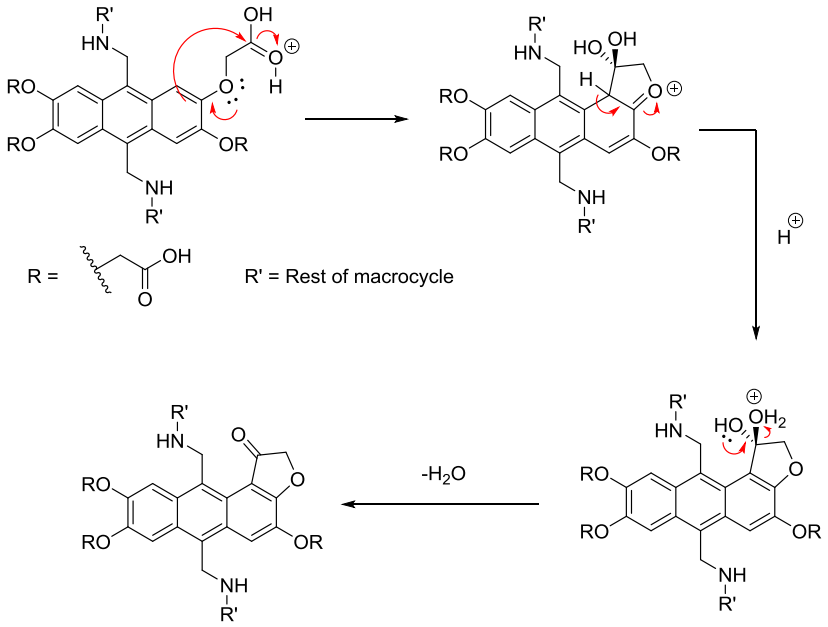
A notable observation during the NMR analysis of this unknown product was that the carbon signal for the anthracenyl CH could not be found. Such a signal had been very consistently observed (at ~105 ppm) throughout the synthetic route *via* ^{1}H-^{13}C HSQC and ^{13}C characterisation. However, this

characteristic signal was not observed during characterisation of the deprotection product. It was also found that despite varying the reaction times from 30 minutes to 16 hours, the product obtained was very consistent which implied that this unexpected reaction was very fast.

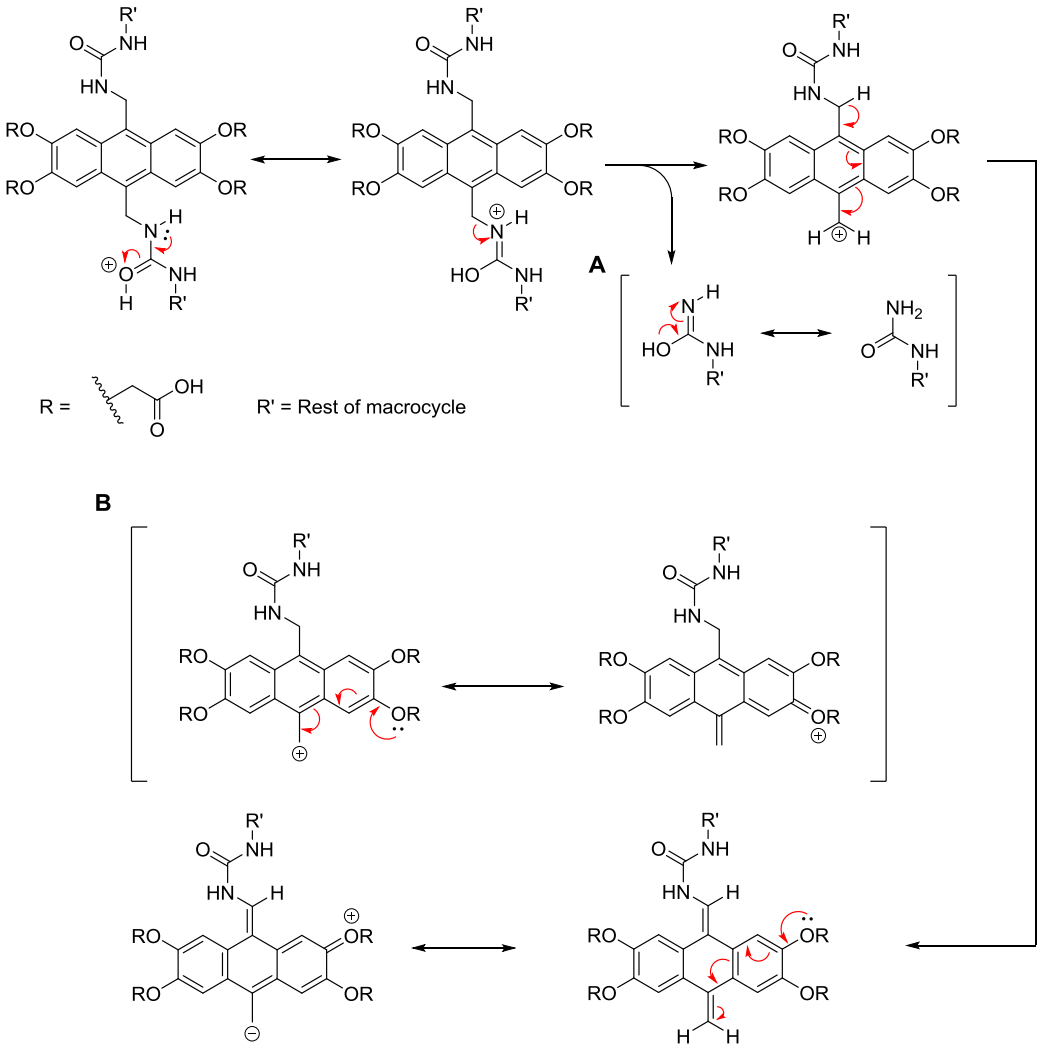
This information would suggest that an acid promoted loss or destruction of the anthracene core was taking place. Three potential acid promoted side reactions were hypothesised. Firstly, α-elimination of water from the carboxylic acid can give an acylium ion (see Hayashi rearrangement and Schmidt reaction).[55,56] This was then followed by a 5-*exo*-*dig* cyclisation and loss of a proton to restore aromaticity which gives a cyclic ketone (Scheme 16). An alternate pathway would be protonation of the carboxyl carbonyl followed by a 5-*exo*-*trig* cyclisation to form a geminal diol, this intermediate can then lose water to form a cyclic ketone (Scheme 17). Both of these reaction pathways would be promoted by the electron donating alkoxy groups on the anthracene facilitating the cyclisation step. It was unknown how many cyclisations could occur per anthracene but this outcome could lead to asymmetrical products – resulting in complicated product NMR spectra. An additional possible side reaction was the acid promoted heterolytic cleavage of the urea at the benzylic position (Scheme 18). This would yield a benzylic cation, which would be stabilised through resonance of electrons in the electron rich anthracene. Loss of a proton can then facilitate formation of a bis-alkene. Loss of aromaticity is compensated by delocalisation of electrons into the anthracene ring system, from the alkoxy substituents. Such a reactive intermediate as the bis-alkene could then react further, leading to a variety of potential products – which would also explain the complex NMR spectra obtained.



Scheme 16 Hypothesised mechanism for the 5-*exo-dig* cyclisation of anthracene and the pendant solubilising chain, *via* an acylium ion intermediate, to give a cyclic ketone.



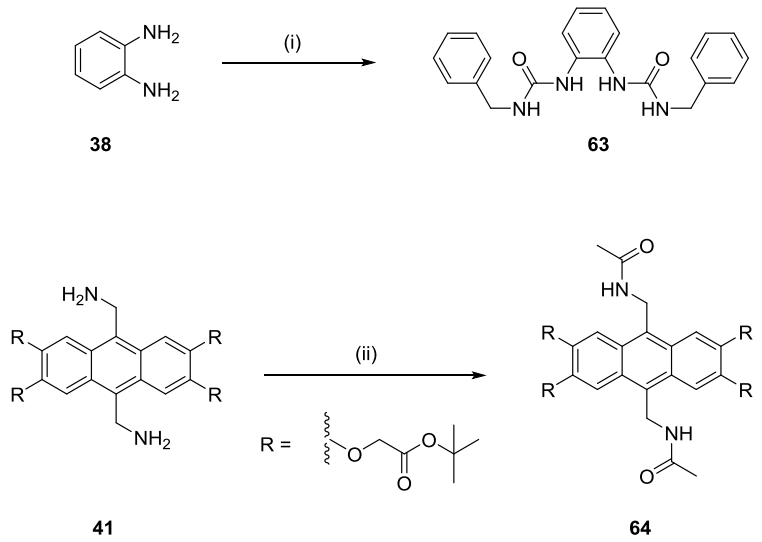
Scheme 17 Hypothesised mechanism for the 5-*exo-trig* cyclisation of anthracene and the protonated carboxylic acid side chain to form a geminal diol. Loss of a proton restores aromaticity, followed by loss of water to yield a cyclic ketone.



Scheme 18 Hypothesised mechanism for the acid promoted heterolytic cleavage of the benzylic ureas **(A)**, generating a resonance stabilised benzylic cation **(B)**. Loss of a proton generates a bis-alkene which is stabilised through resonance of the alkoxy groups.

To elucidate which undesired reaction was taking place, as well as probe the instability of the macrocycle to acidic conditions, two model substrates were synthesised. One of these was designed to represent the urea spacers **63** and the other to represent the anthracene surface **64** of the macrocycle (Scheme 19). Model bis-urea **63** was synthesised and then exposed to TFA (50% v/v) in dichloromethane, with NMR of the crude reaction mixture showing no change after 2 days. Acetylation of diamine **41** afforded **64** which was then stirred in TFA (50% v/v) and dichloromethane for 16 hours. NMR analysis of the product after neutralisation to pH 7 revealed a mixture of products had been obtained. Some signals corresponded to those seen in the attempted deprotection of macrocycle **37**, but the spectra obtained were still highly complex so full characterisation was not possible. These results would initially suggest that the anthracene unit was the source of the problem – possibly due

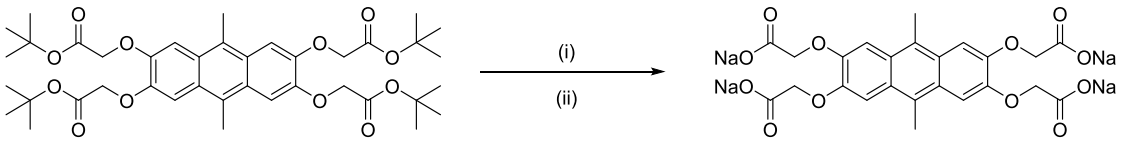
to the intramolecular cyclisations proposed in Scheme 16 and Scheme 17. The stability of **63** to acid infers that the heterolytic cleavage of the ureas (Scheme 18) was less likely. However, there was no reason to suggest such a reaction was not occurring for **64**, since the anthracene units stabilise the benzylic cation much more effectively than the simple benzene units of **63**.



Scheme 19 (i) benzyl isocyanate, CH_2Cl_2, reflux, 16h, 90%; (ii) Acetyl chloride, DIPEA, CH_2Cl_2, rt, 16h, 84%.

In attempts to confirm which side reaction was occurring to macrocycle **37** and test substrate **64**, further test reactions were performed on intermediate **44** (Table 2). Tetra ester **44** was subjected to a range of TFA concentrations for a set reaction time (16 hours), then neutralised to pH 7 and filtered, with the yield recorded and purity of product observed by ^{1}H NMR. It was found that varying the concentration of TFA had no effect on the purity of the product obtained, with all test reactions yielding pure tetra carboxylate **65** (by ^{1}H NMR analysis in D_2O). Varying the concentration of TFA did have a dramatic effect on the yield of reaction however, with full conversion of **44** to **65** only observed when >30% (v/v) of TFA was used. It was speculated that lower concentrations of TFA reduce the rate of reaction, resulting in incomplete deprotection of **44**. Such partially deprotected intermediates would not be soluble in water at pH 7, and are thus likely filtered out during the neutralisation process. This would explain the lower yields but pure products obtained for the lower concentrations of TFA. Most curiously however, was the complete lack of degradation of the anthracene when exposing **44** to high concentrations of TFA. This lends to the theory that no intramolecular cyclisations are taking place during deprotection of macrocycle **37** (Scheme 16 and Scheme 17) and that the decomposition observed was largely due to the cleavage of the ureas (Scheme 18) *via* a benzylic cation (stabilised by the electron rich anthracene).

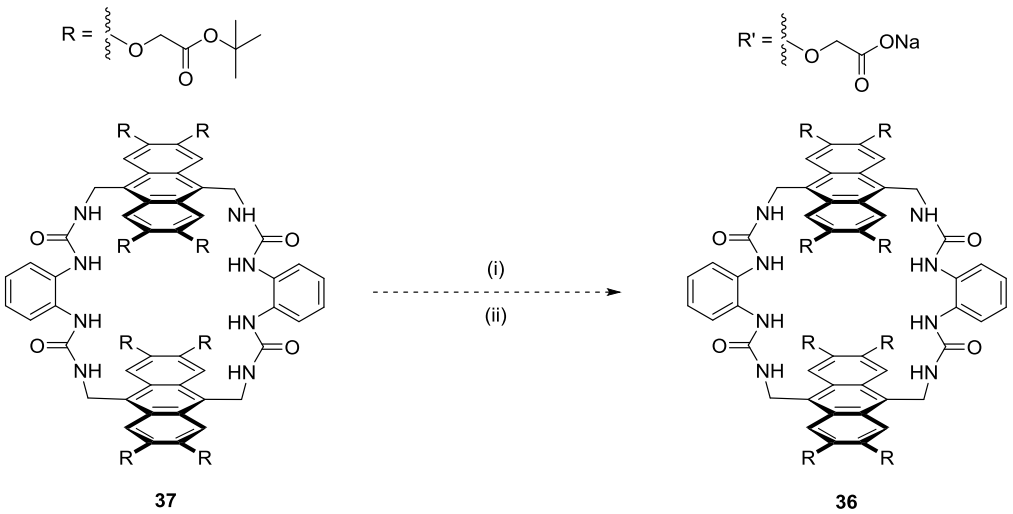
Table 2 TFA concentrations screened for test substrate **44**; (i) TFA, Et_3SiH, CH_2Cl_2, 16h, rt (ii) NaOH (aq), H_2O.



44 **65**

Entry	TFA (% of total volume)	Yield (%)	Comments
1	1	10	Pure after neutralisation
2	5	21	Pure after neutralisation
3	10	32	Pure after neutralisation
4	20	67	Pure after neutralisation
5	30	96	Pure after neutralisation
6	50	Quantitative	Pure after neutralisation.

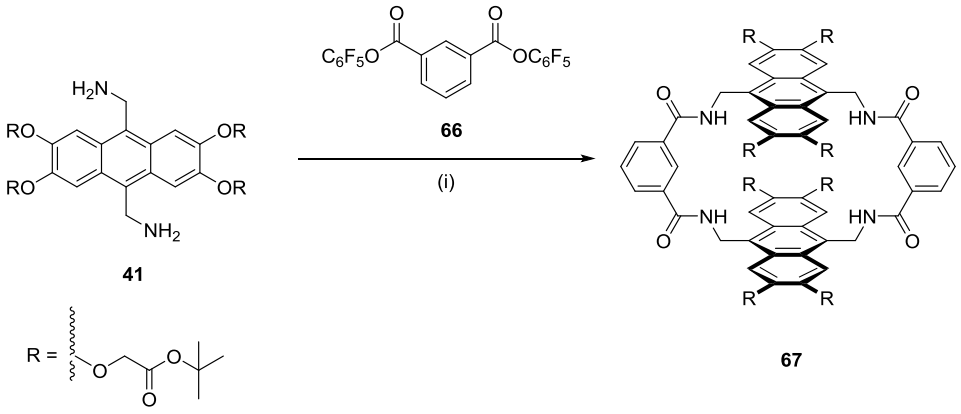
Despite this relationship of concentration of acid and yield of tetra carboxylate **65**, it was decided to attempt the deprotection of macrocycle **37** with much lower concentrations of TFA (i.e. <10% v/v). It was speculated that a potential balance could be reached, where the concentration of TFA would be enough to drive the reaction to completion without decomposing the product. Macrocycle **37** was then exposed to several concentrations of TFA (1%, 2% and 5% v/v) for 6 hours, and the crude product analysed by ^1H NMR (in CD_3OD) before neutralisation to pH 7. Signals characteristic of the *tert*-butyl groups were clearly visible in the ^1H NMR and confirmed by observation of the $\underline{C}(CH_3)_3$ quaternary sp^3 carbon using ^1H-^{13}C HMBC. No degradation was observed which was deemed promising. However the reaction was now very slow to proceed, due to the decreased amount of TFA. Disappointingly, increasing the concentration of TFA to 10% (v/v) and above appeared to cause degradation of macrocycle **37**, while still not removing all the *tert*-butyl groups. This would suggest the proposed urea cleavage occurs at a similar or faster rate than the deprotection, meaning the current deprotection strategy and conditions are not workable. No purification was deemed possible at this stage, so even if fully deprotected macrocycle **36** had formed there would be no way of accessing it. Later work revealed that the deprotected macrocycles could be purified by reverse phase HPLC (see Chapter 4, Figure 37), but this was not discovered in time to be applied to this work.



Scheme 20 (i) TFA (1-5%), Et_3SiH, CH_2Cl_2, rt 6h; (ii) NaOH (aq), H_2O.

Macrocycle **67** (Scheme 21) was now proposed as a model substrate to investigate how sensitive the deprotection conditions were to changes in the linker functionality – with **67** employing a bis-amide spacer unit akin to more classical receptors such as anthracene receptor **25**. This would demonstrate whether the electronics of the anthracene unit stabilising the benzylic cation are the source of the issues, and not the bis-urea linker functionality. If **67** did successfully deprotect to give a clean product, this would still yield a novel receptor and interesting comparison to **25**.[35] The effects of removing the dendrimer groups on binding properties has not been investigated before and this new anthracene surface may yet yield new interesting activities towards guest molecules. Synthesis of **67** proceeded smoothly. Reaction of diamine **41** with bis-pentafluorophenyl ester **66** gave **67** in 26% yield after purification by reverse phase HPLC. However, as seen before for **37**, macrocycle **67** also displayed rapid oxidation of the anthracene surfaces in organic solvents (see Scheme 13). Oxidation would routinely occur during NMR characterisation – especially during the experiments with longer acquisition times (i.e. 1D ^{13}C spectra, Figure 26). Attempts to remove oxygen from NMR solvents and use of amberised glass NMR tubes appeared to have little improvement. It is possible that the macrocycle resides in a conformation that promotes this oxidation, even more so than for **37**. Use of protic NMR solvents (such as CD_3OD) to suppress oxidation were not viable due to insolubility of **67** and so full ^{13}C characterisation of **67** was not achieved. Despite incomplete NMR characterisation, **67** was exposed to similar acidic deprotection conditions as for **37**. Neutralisation to pH 7 and subsequent ^1H NMR analysis (in D_2O) yielded a similarly complex spectrum as seen when attempting to deprotect **37**. While the spectrum was not entirely consistent with the deprotection products seen for **37**, differences were rationalised to be due to the different spacer units used for **67**. It was very evident however that a single

symmetrical product had not been obtained. This result was deemed to corroborate with previous hypotheses: that the tetra carboxyl anthracene unit was the source of the issues by promoting cleavage of the benzylic nitrogen bonds to form a stabilised benzylic cation.



Scheme 21 (i) DIPEA, CH_2Cl_2, rt, 3 days, 26%.

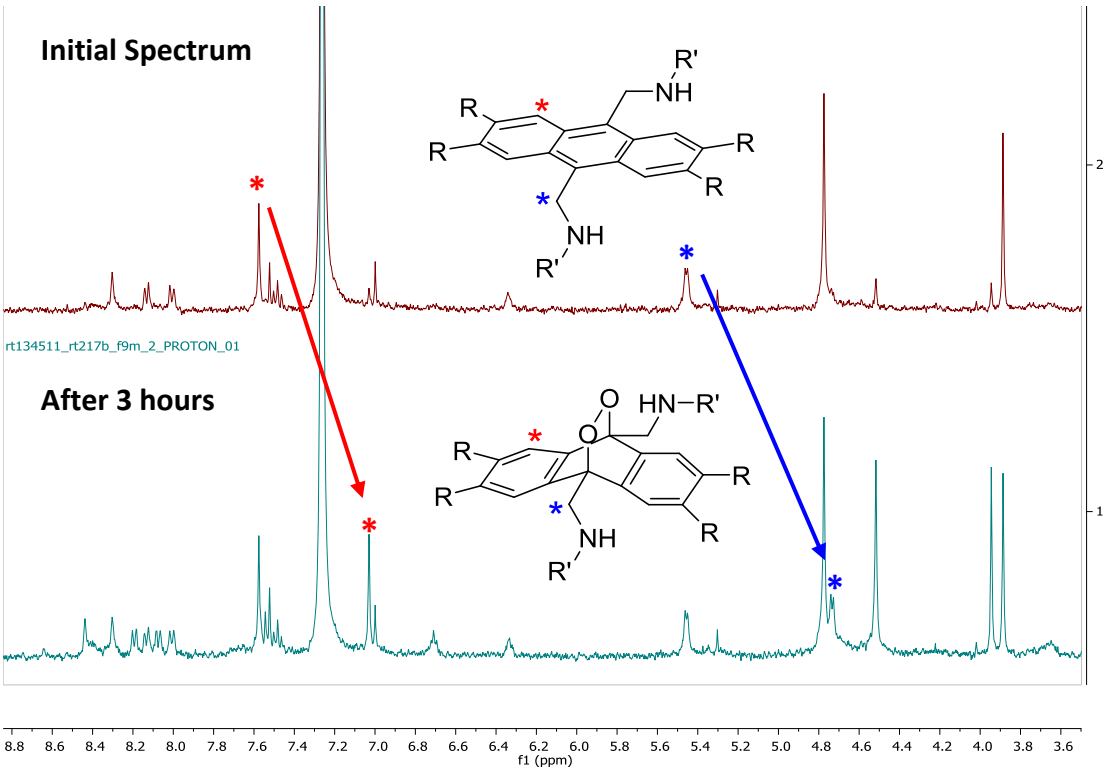


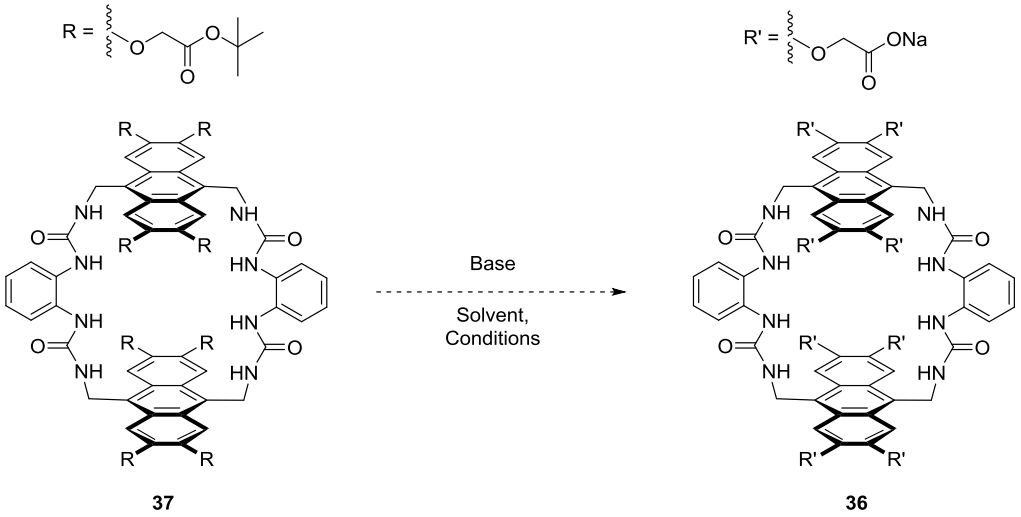
Figure 26 Partial ^{1}H NMR (500 MHz, $CDCl_3$, 298 K) of isolated receptor **67** (top) and partially oxidised receptor (bottom) after 3 hours in deuterated chloroform. Characteristic changes in chemical shift upon oxidation are highlighted.

Formic and acetic acid were then screened to see if weaker acids would reduce the amount of decomposition of **37**. However, very little to no deprotection was observed when using weaker acids – even at high concentrations (>50% v/v of acid used). Varying the solvent was also investigated, with the use of more polar solvents lowering the acidity of TFA and in turn prevent the hypothesised cleavage of the benzylic ureas. However, as seen for the weaker acids, no reaction was observed when using TFA in alcohols. A reaction system of TFA in methanol was attempted to deprotect the *tert*-butyl esters and then convert them to the equivalent methyl esters, which could then be hydrolysed under mild conditions with aqueous base. Again, no deprotection of the *tert*-butyl groups was evident.

There is literature precedent for base mediated hydrolysis of *tert*-butyl esters to the carboxylate but under forcing conditions with long reactions times.[57] Despite this, various basic conditions were screened for the deprotection of **37** (Table 3). Unfortunately, none of the desired macrocycle **36** was observed by ^{1}H NMR (in D_2O) after neutralisation to pH 7 with acidic ion exchange resin. The NMR spectra obtained after using these conditions were arguably more convoluted than those seen for the acidic deprotections – a possible indicator for multiple species. The concentration of sodium hydroxide, as well as reaction time, were screened (entries 1-3, Table 3) to probe whether the 'milder conditions' (i.e. entry 1) were not enough to enable the reaction to reach completion. However, all produced complex mixtures.

Entries 4 and 5 were derived from procedures for the hydrolysis of *tert*-butyl esters at room temperature using a 1:1 ratio of THF and methanol. However, the reaction at room temperature (entry 4) appeared to be very slow as the NMR showed several *tert*-butyl ester species, presumably due to incomplete reaction. Refluxing this procedure (entry 5) instead appeared to cause a loss of urea linker signals, as did entry 6 when using tBuOK. Previous studies demonstrate that the hydrolysis of substituted ureas is possible under strongly basic conditions, which might explain the complex spectra obtained using these procedures. However, when model bis-urea **63** was refluxed in strongly basic conditions for 3 days (same as entry 3) there were no signs of hydrolysis. It could be speculated that macrocycle **37** is more sensitive to urea hydrolysis, particularly if the system resides in a higher energy strained conformation - as cleavage of the ureas would favourably relieve this strain. The long reaction times and forcing conditions could promote oxidation of the anthracene as well, even though all efforts were maintained to eliminate oxygen from the system.

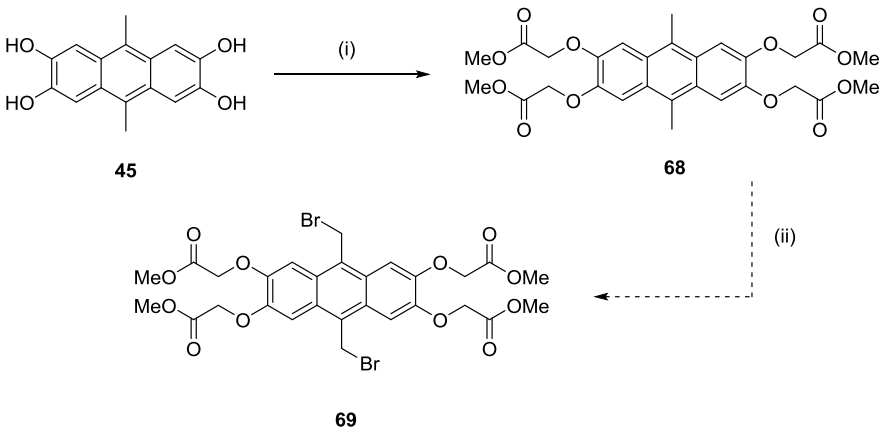
Table 3 Bases, solvents and conditions screened for the base mediated hydrolysis of **37**.



Entry	Base	Solvent	Conditions
1	aq. NaOH (3M)	THF	reflux, 1 day
2	aq. NaOH (4M)	THF	reflux, 1 day
3	aq. NaOH (6M)	THF	reflux, 2 days
4	aq. NaOH (6M)	THF:MeOH (1:1)	rt, 1 day
5	aq. NaOH (6M)	THF:MeOH (1:1)	reflux, 2 days
6	tBuOK	THF	rt, 1 day

Evidently, the *tert*-butyl ester was not a suitable protecting group for macrocycle **37** given the observed sensitivity to acid. Therefore other routes to the carboxylate were considered. The simplest modification was to replacewitch the *tert*-butyl esters for methyl esters, which were expected to be sufficiently susceptible to hydrolysis under mild basic conditions. It was envisaged that the synthetic route would remain largely the same. Indeed, synthesis of tetra methyl ester **68** proceeded smoothly from tetra-ol **45** using methyl bromoacetate (Scheme 22). Conversion to the bromide **69** was found to be not possible however. Following work up and extraction, analysis of the crude material by ^{1}H NMR gave a complex mixture featuring multiple methyl ester species. Signals relating to the anthracene were severely reduced, as was mass recovery of the product which was no longer fluorescent. Larger atoms such as bromides and iodides are known to quench fluorescence through collisional quenching

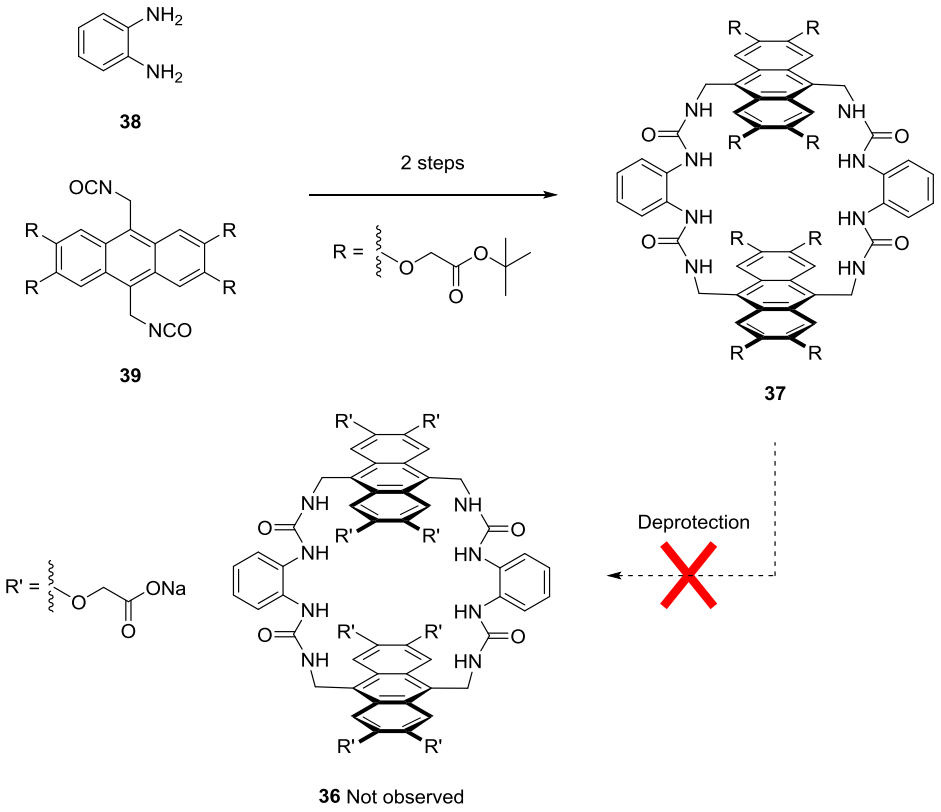
mechanisms which could explain this observation.[58] The aqueous extraction phase was observed to be highly fluorescent however and it was initially postulated that the basic aqueous work up could be the source of the problem through potential hydrolysis of the methyl esters. However, analysis of the crude reaction mixture before work up also gave a complex mixture, with no major product discernible, suggesting the aqueous work up was not degrading the product. It was then hypothesised that the radical reaction was somehow interfering with the methyl ester side chains, although by what mechanism this could occur is unknown at this time. Considering the equivalent reaction with *tert*-butyl esters occurs in quantitative yield, it would seem rational to suggest the newly substituted methyl esters are responsible for the failed reaction here and this synthetic route is impractical.



Scheme 22 (i) methyl bromoacetate, K_2CO_3, THF, reflux, 24h, 65%; (ii) NBS, ABCN, CH_2Cl_2, reflux, 1h, bromide **69** not observed.

Other protecting groups strategies were considered after this, such as silicon containing esters that could be deprotected under very mild conditions with fluoride.[59] However, after the numerous issues encountered with using highly substituted electron rich anthracene units (including acid instability and increased risk of oxidation), it was eventually decided to focus on a different receptor design that would not be susceptible to such problems.

2.5 Conclusions



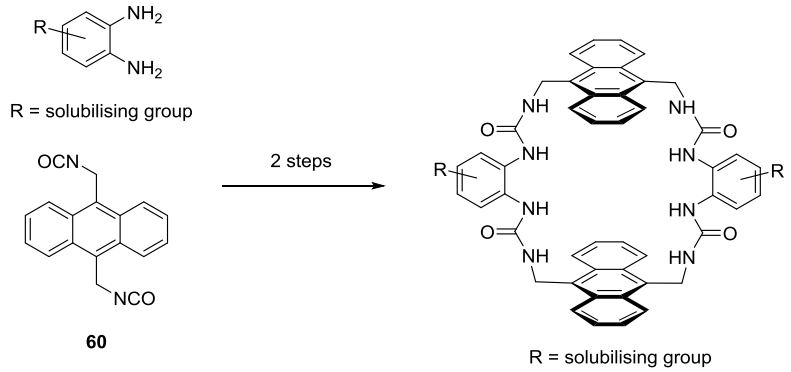
Scheme 23 Successful methodology (Scheme 12 and Scheme 14) developed to access tetraurea macrocycle **37** from diamine **38** and bis-isocyanate **39**. Deprotection to water soluble macrocycle **36** was not achieved.

Synthetic methodology towards incorporation of bis-urea spacer units into macrocycles was achieved, with the successful synthesis of tetraurea **37** after optimisation of the synthetic route (Scheme 23). Subsequent attempts to access the water soluble macrocycle **36** were not successful however (Scheme 23). The deprotection of the *tert*-butyl groups was probed using reaction condition screening, in depth NMR studies and synthesis of test substrates. This work revealed that the electron rich anthracene units were potentially facilitating cleavage of the benzylic ureas *via* a benzylic cation intermediate, when exposed to acidic conditions. The electron rich anthracenes were also extremely susceptible to oxidation, making the macrocycles and intermediates difficult handle. No solution to obtaining **36** was able to be determined at the time of this work however, and it was concluded that heavily substituted electron rich anthracene units are not viable substrates to work with due to their instability to acid and sensitivity to oxidation. Future efforts would be directed towards other receptor designs that would not suffer from these problems.

Chapter 3 – Development of a Diamino Spacer Unit

3.1 Design Rationale and Potential for Future Receptor Designs

Following on from the successful synthesis of anthracene macrocycle **37** (and problems faced regarding instability of the macrocycle), efforts were directed towards installing a solubilising group on the urea spacer unit. To do so would require synthesis of an appropriate 1,2-diaminobenzene 'linker' with the solubilising group pre-installed. Development of this linker unit would enable the use of a much simpler anthracene core, when compared to the electron rich anthracene used for **37** (Scheme 24). This was predicted to circumvent the previous issues of instability to acid and oxygen.



Scheme 24 Proposed receptor design, featuring a new diamino linker unit and simple anthracene.

Several potential designs could be envisaged, with maintaining symmetry of the macrocycle being a key consideration (**Figure 27**). Other desirable attributes would be a rigid ring structure to assist in maintaining a cage conformation of the macrocycle, as well as electron withdrawing groups to amplify the hydrogen bonding character of the ureas in the target receptor. Once a route towards this diamino linker had been established, it was to be incorporated into an anthracene based receptor (Scheme 24) and tested for the binding affinity towards several carbohydrates (as proposed for **36**, Chapter 2).

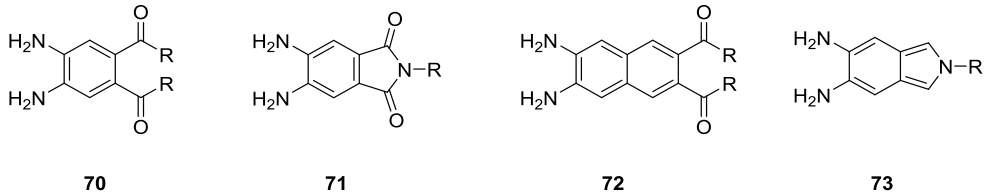


Figure 27 Some examples of possible linker targets, where R = dendritic solubilising group.

Proceeding this, now that the solubilising group was part of the spacer unit, various anthracene derivatives that alter the optical and electronic properties towards binding were to be investigated. Modularity would also exist with regards to the solubilising group, which could be tuned to optimise solubility, functionality at various pH ranges or even facilitate attachment to a solid state support or surface. This new linker could even be incorporated into different aromatic surfaces, thus generating a new library of receptors. The Davis group and others have developed methodology towards various aromatic amines previously (Figure 28).[33,35,38,39,60] Conversion of these amines to the corresponding isocyanate and reaction with this new linker would yield the equivalent urea receptors – comparisons of which to the previous amide receptors would be most interesting.

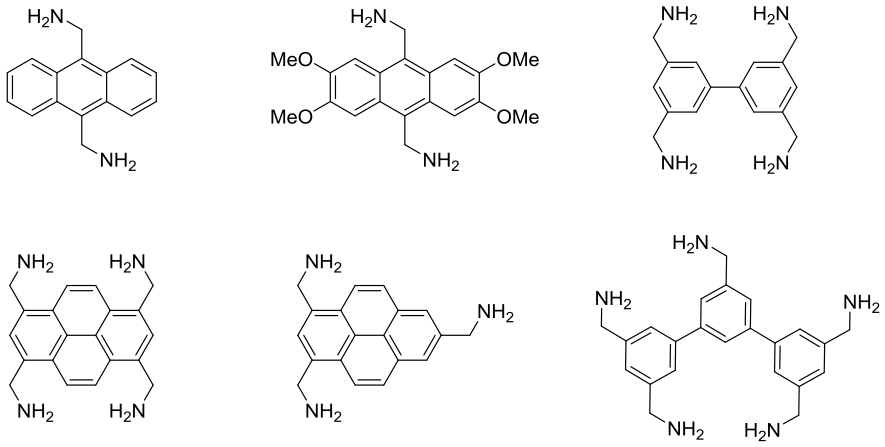
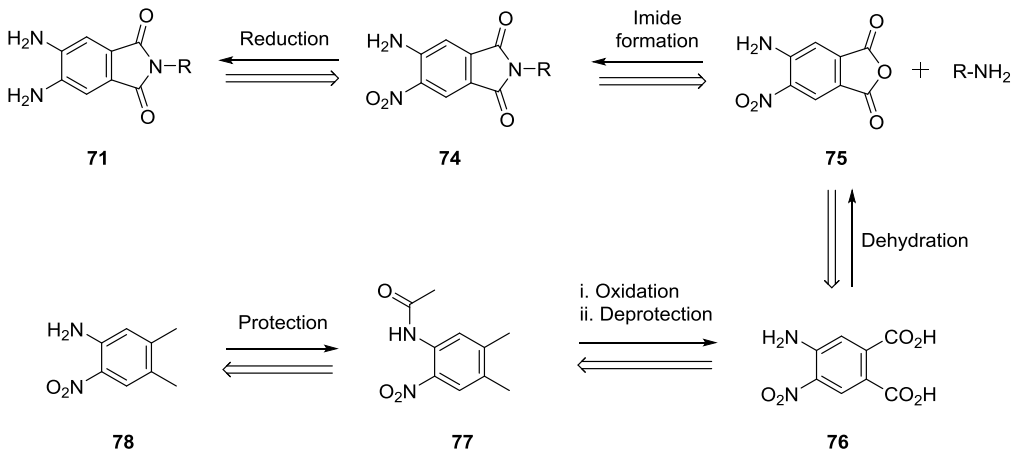


Figure 28 Examples of aromatic amines previously synthesised by the Davis group that have been incorporated into carbohydrate receptors.[33,35,38,39,60]

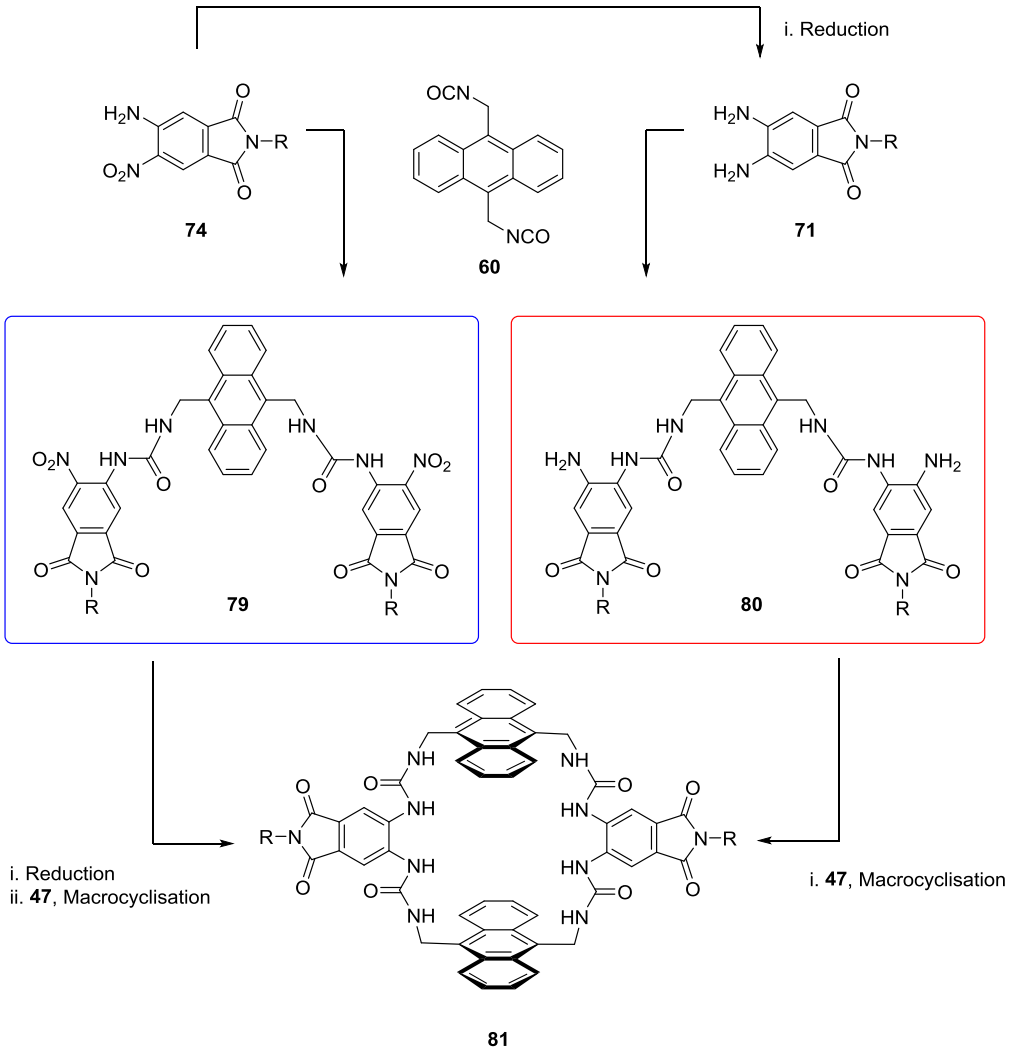
3.2 Synthesis of Diamino Phthalimide Spacer

The initial diamino spacer target to be pursued was diamino phthalimide **71**, as this featured all the desired properties of symmetry, rigidity, and electron deficiency, as well as being synthetically accessible. The route towards **71** had already been partly undertaken in previous literature, which made the synthesis all the more theoretically feasible.[61] Retrosynthetic analysis revealed that **71** could be accessed in 5 steps from commercially available 2-nitro-4,5-dimethyl aniline **78** (Scheme 25). The key step being imide formation from anhydride **75** and a suitable amine that features a pendant solubilising group.

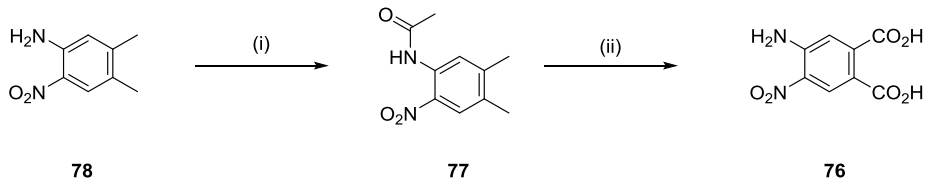


Scheme 25 Retrosynthetic analysis of the proposed diamino phthalimide **71**, where R = solubilising group.

The inherent protection of one of the amines as a nitro group **74** provides potential for two routes to the desired macrocycle (Scheme 26). Reduction to the diamine **71** and then reaction with diisocyanato anthracene **60** leads to the amino urea half receptor **80** (red box, Scheme 26). This intermediate can then react with another equivalent of isocyanate to yield the tetraurea macrocycle **81** in a similar fashion to **37** seen before (see Chapter 2). However, reaction of the unreduced nitro aniline **74** would afford the nitro urea half receptor **79** (blue box, Scheme 19). This could be a potentially cleaner reaction when compared to reaction with diamine **71**, as there would be no potential for further reaction of the half receptor with more isocyanate. Reduction to the amine and subsequent cyclisation would then give the target macrocycle **81**. Both routes were to be investigated for viability during the synthesis and were predicted to be largely reliant on how facile the reduction of the nitro groups of both **74** and **79** will be, as well ease of formation of both half receptors **79** and **80**.



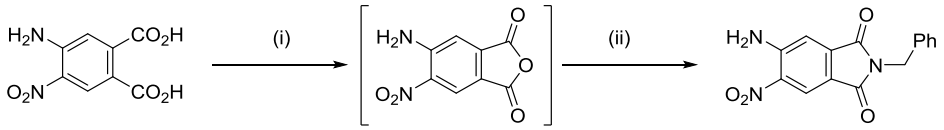
Scheme 26 Two possible routes to the target macrocycle **81** *via* two different half receptor intermediates, where R = solubilising group.



Scheme 27 (i) Acetic anhydride, Acetic Acid, reflux, 2h, 98%; (ii) $KMnO_4$, H_2O, reflux, 3 days, 67%.

Commercially available 2-nitro-4,5-dimethyl aniline **78** was acetylated with acetic anhydride in refluxing acetic acid to afford **77** in near quantitative yield (Scheme 27). Acetylated amine **77** had been oxidised to the bis-acid **76** previously in the literature using potassium permanganate, although no reaction temperature or time was given in the procedure.[61] It was found that oxidation with potassium permanganate in refluxing water for 3 days gave **76** in good yield, with no purification necessary after work up. It was also found that the oxidative conditions unexpectedly resulted in deprotection of the

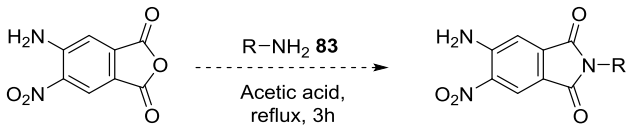
acetyl group as well, effectively removing the step from the route. Acetylation of **78** was found to be critical to oxidation, as attempted oxidation of **78** using the same conditions, resulted in near quantitative recovery of starting material. Esterification with methanol and catalytic H_2SO_4 allowed purification and subsequent characterisation of the reaction products. It was found that the main product was the desired bis-acid **76** and the minor products were two mono carboxylic acid isomers. Therefore, future procedures featured additional equivalents of potassium permanganate being added part way through the reaction. This resulted in a reduction of yield of the mono carboxylic acid isomers, giving an optimised yield of 67% for **76**.



76 **75** **82**

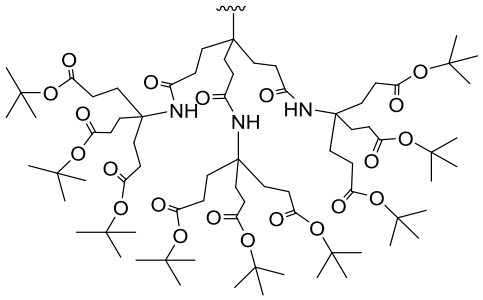
Scheme 28 (i) Acetic anhydride, THF, reflux, 16h; (ii) benzylamine, acetic acid, reflux, 3h, 87%.

Dehydration of **76** to the anhydride **75** was achieved by addition of acetic anhydride and then refluxing in THF (Scheme 28). Conversion to **75** was inferred by TLC and ^{1}H NMR of the crude reaction mixture, with **75** being used immediately without purification in the next step to form the phthalimide. To test phthalimide formation, benzylamine was added to **75** and then refluxed in acetic acid to promote ring closure to benzyl phthalimide **82**, which was isolated in an excellent yield of 87% after column chromatography.



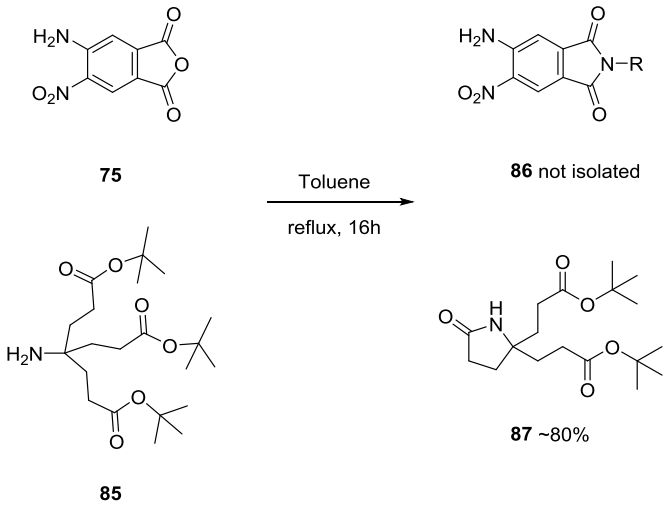
75 **84**

R =

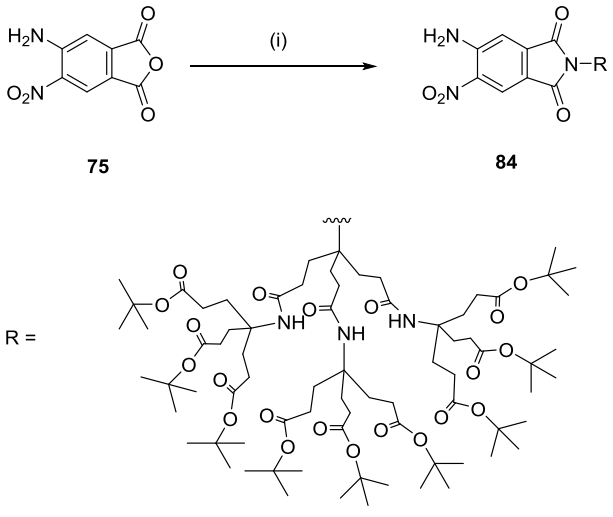


Scheme 29 Initial attempts to form phthalimide **74** from the 2nd generation solubilising group amine (G2MM) **83**.

With model phthalimide **82** successfully synthesised, the procedure was repeated with the 2nd generation amine solubilising group (G2MM, Scheme 29) **83**, which was prepared according to previous literature.[36] However, none of the desired product was observed in the crude reaction mixture by mass spectrometry or after attempted purification. It was thought that the *tert*-butyl esters of **74** might be partially deprotected in the acidic solvent, and that the high temperatures help force this deprotection even in relatively weak acid (compared to TFA). The reaction solvent was then changed to toluene and the test reaction repeated to form model benzyl phthalimide **82**, which was again synthesised in excellent yield (84%). To test this new procedure, the more synthetically accessible 1st generation amine solubilising group (G1M) **85** was used, which was prepared according to previous literature.[36] Upon purification however, none of **86** was isolated. It was in fact discovered that the majority of **85** had cyclised to lactam **87** (Scheme 30). This unwanted reaction has been previously reported to occur in polar solvents at high temperatures but was not known to occur in more apolar media until now.[62]

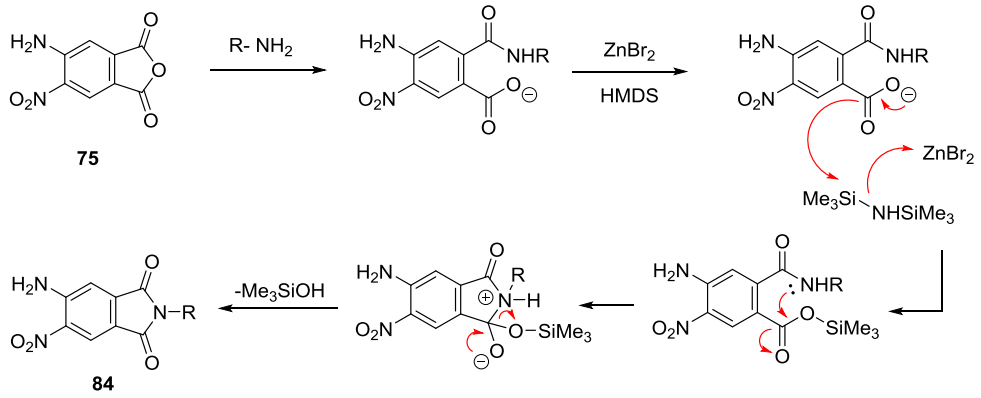


Scheme 30 Attempted phthalimide formation using the 1st generation amine solubilising group (G1M).



Scheme 31 (i) R-NH_2 **83**, toluene, 60 °C, 30 min, then $ZnBr_2$, HMDS, 60 °C, 16h, 64%.

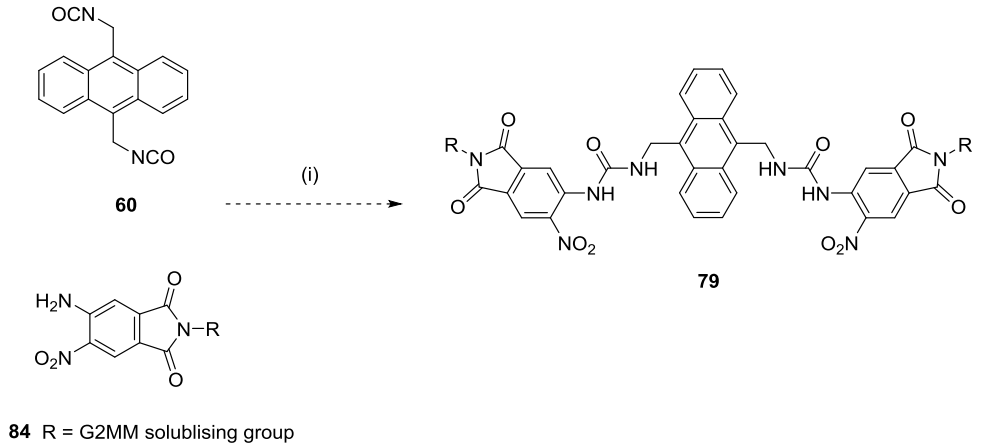
This procedure was then repeated with G2MM amine **83**, which cannot cyclise to form a lactam unlike G1M **85**. None of Phthalimide **84** was isolated initially, although some consumption of starting materials did occur – an indication that amine **83** had reacted with the anhydride but would not cyclise to form the imide. It was only when zinc bromide and HMDS were added after 30 minutes of initial reaction that the reaction appeared to proceed further, with the zinc bromide acting as a Lewis acid to facilitate formation of the silicon ester intermediate which can then promote the final ring closure step to afford the imide (Scheme 32).[63] Phthalimide **84** was then isolated in a good yield after column chromatography. It was rationalised that the reaction was not as high yielding as for **82** due to the bulky nature of **83** reducing the rate of attack at the anhydride or rate of cyclisation. This may allow more time for the anhydride or silicon esters to hydrolyse, evidence for this was that some amine (~20%) was isolated during purification – suggesting that not all amine reacted with the anhydride.



Scheme 32 Proposed mechanism for the formation of the phthalimide **84** *via* a silicon ester intermediate.

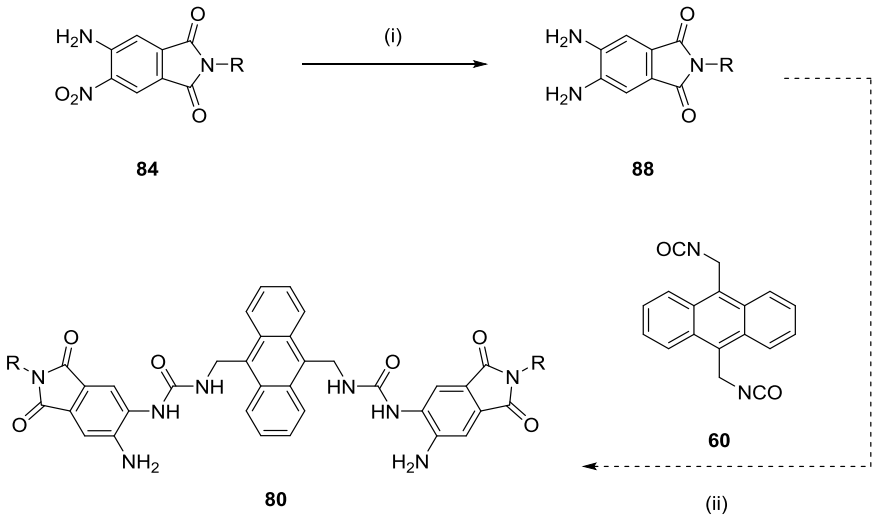
Phthalimides **82** and **84** were then tested for reaction with anthracene diisocyanate **60** to form half receptor **79** (Scheme 33). However, it was found that no reaction was observed upon stirring either **82** or **84** with **60** in dichloromethane at room temperature or reflux. It was thought that both the electron withdrawing nitro and imide groups render the amine a very poor nucleophile. Addition of organic bases such as Et_3N and DIPEA appeared to have no effect. Addition of relatively strong organic base DBU to a solution of **84** and **60** in dichloromethane induced a distinct colour change from yellow to red/orange – believed to be due to deprotonation of the amine. A very fast reaction was then observed when monitoring with TLC, with the isocyanate rapidly consumed (<1 hour). However, none of **79** was isolated after purification with column chromatography. Upon this result, it was then discovered that DBU has been previously reported to undergo adduct formation with isocyanates to form thermally stable heterocyclic structures for use as polymerisation initiators.[64] With this information in hand, the reaction was repeated but with a slow addition of isocyanate **60** to a solution of **84** and DBU. It was speculated that this method would minimise interaction of free basic DBU and isocyanate. The

[M+2Na]$^+$ adduct of the product was observed using ESI mass spectrometry of the crude reaction mixture, but no product was isolated upon purification. This may be due to **84** more easily forming charged adducts under mass spectrometry conditions compared to the other reaction components and that only very small amounts of **84** are forming in the reaction. Cooling the reaction to 0 °C did not appear to prevent this unwanted reaction either.



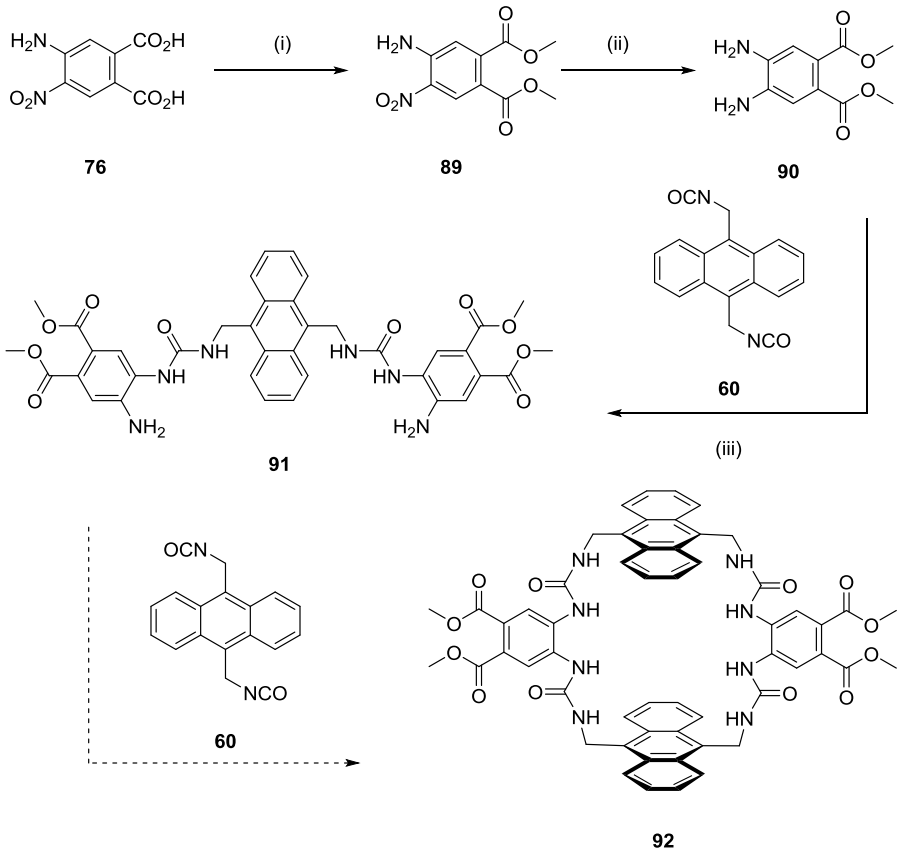
Scheme 33 (i) DBU, CH_2Cl_2, rt, 1h, **88** not isolated.

Reduction of the nitro group of **84** would afford the diamino phthalimide, increasing the nucleophilicity of the amines towards isocyanate **60**. Diamino phthalimide **88** was synthesised in excellent yield after reduction of **84** using Pd/C and 1 bar pressure of hydrogen (Scheme 34). Attempts to form half receptor **80** from phthalimide **88** and diisocyanate **60** were unsuccessful however. No reaction was observed at room temperature or when refluxed in dichloromethane (Scheme 34). As seen for nitro phthalimide **84**, addition of organic bases also appeared to have very little positive effect. Synthesis of previous half receptor **59** (Chapter 2) used vast excesses of diamine to drive the reaction (up to 20 equivalents) – something not possible here to due to limited quantities of **88**. It is possible that the amines may still be deactivated *via* the electron withdrawing effects of the imide, or the bulky solubilising group hinders approach to the isocyanate. Initially, further investigations into this lack of reactivity were suspended and other synthetic ventures pursued.



Scheme 34 (i) H_2, Pd/C, MeOH, rt, 3h, 90%; (ii) CH_2Cl_2, reflux, 48h, **80** not observed. R = G2MM solubilising group

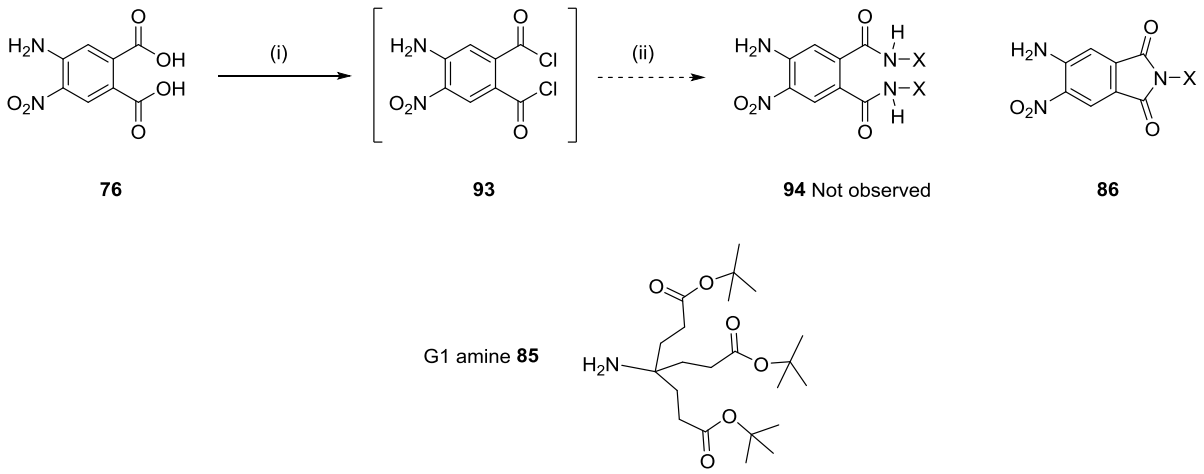
Future work covered in Chapter 5 found that addition of acylation catalysts such as pyridine and DMAP greatly increased reactivity of electron deficient anilines towards isocyanates. This was then applied in a limited capacity to some intermediates discussed earlier in this chapter. Nitro bis-acid **76** was converted to bis-methyl ester **89** *via* Fischer esterification in methanol and then reduced to diamine **90** with hydrogen and Pd/C in excellent yield (Scheme 35). Reaction of this diamine **90** with anthracene isocyanate **60**, in the presence of DMAP, afforded half receptor **91** in good yield. Half receptor **91** proved relatively insoluble in most organic solvents, but this allowed straightforward purification by filtration. Limited NMR characterisation was carried out for **91**, but the product was confirmed by high resolution ESI mass spectrometry – where the $[M+2H]^{2+}$ ion was observed. While this methodology could be applied to the earlier phthalimide systems (Scheme 34), there was not sufficient time or material to attempt this, although it is expected that addition of DMAP would allow access to half receptor **80**. Intermediate **91** is arguably more versatile however, as functional group conversion of the methyl esters could afford several derivatives including phthalimides and bis-amides featuring large solubilising groups (such as G2MM). Future work will therefore focus on development of **90** as a linker unit, to generate methyl ester receptors (such as **92**, Scheme 35) which can then undergo late stage derivatisation to generate libraries of receptors more quickly.



Scheme 35 (i) H_2SO_4, MeOH, reflux, 3h, 80%; (ii) H_2, Pd/C, MeOH, 1h, 93%; (ii) **60**, DMAP, CH_2Cl_2, reflux, 16h, 76%.

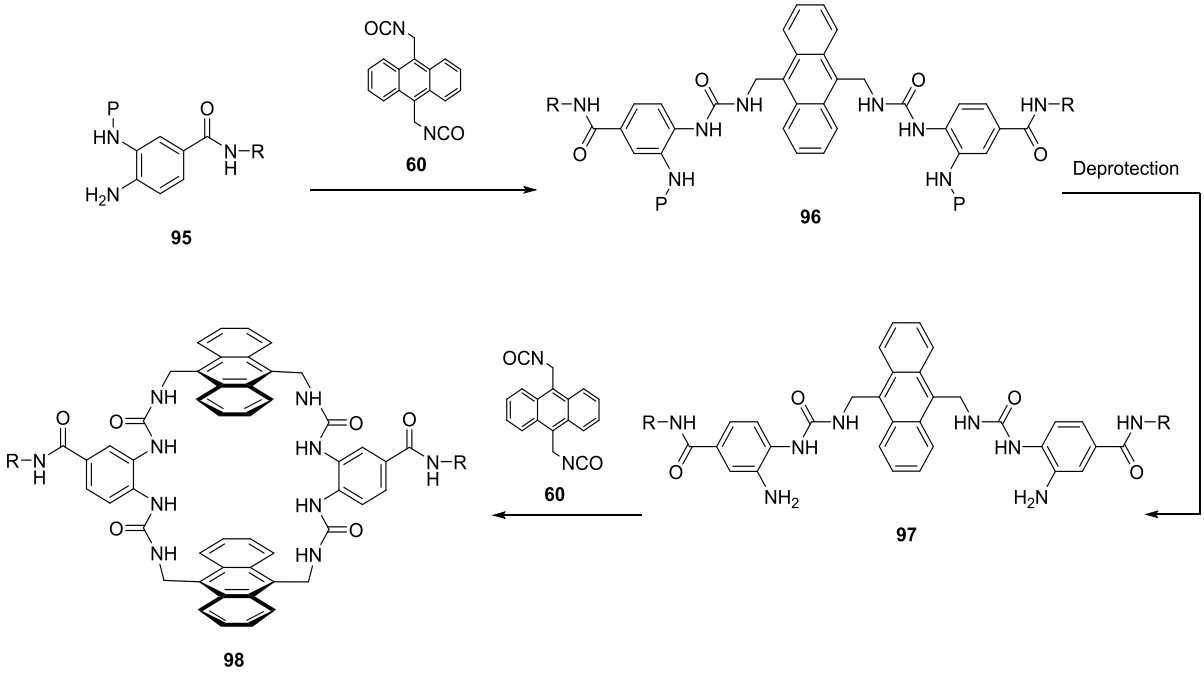
3.3 Synthesis of Diamino Amide Spacer

With the previously synthesised phthalimides initially proving unreactive towards isocyanates, other linker units were targeted such as bis-amide **94**. The route to such compounds had been previously outlined in the literature, and initially appeared particularly suitable as previously synthesised bis-acid **76** could be used.[65,66] Conversion of **76** to the bis-acyl chloride **93** was achieved using catalytic DMF in $SOCl_2$. Formation of **93** was inferred by TLC and ^{1}H NMR of the crude reaction mixture. Acyl chloride **93** was then used without further purification (Scheme 36). Reaction of **93** with G1M amine **85** however did not yield the desired bis-amide **94**. A small amount of phthalimide **86** was isolated upon purification however, suggesting that a cyclisation takes place upon addition of one amine to the acyl chloride to give the imide. Additional equivalents of G1M amine **85** were used to suppress the cyclisation, but this appeared unsuccessful as **94** was not isolated upon purification. It could be speculated that addition of a second amine is comparatively slow due to the bulky nature of the dendrimer, and that the rates of cyclisation and hydrolysis outcompete this. It was then decided that this methodology was not suitable and therefore not pursued any further.

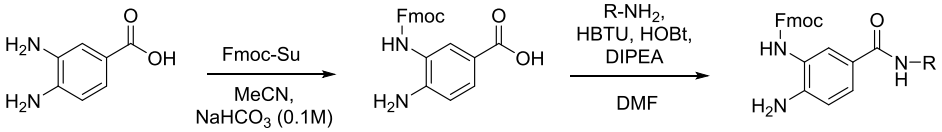


Scheme 36 (i) DMF, $SOCl_2$, (ii) G1M amine **85**, DMAP, Et_3N, CH_2Cl_2, rt, 16 h.

A simpler linker unit was then proposed: mono-amide **95** was postulated to be sufficiently reactive at the amines to form ureas upon addition to an isocyanate (such as **60**). Being unsymmetrical however would mean that one of the amines would require functional group protection before half receptor formation to avoid isomers when forming the macrocycle (Scheme 37). Fortunately, recent literature precedent was set where 3,4-diaminobenzoic acid **99** was mono-Fmoc protected **100** and then able to undergo successful amide coupling to a solid phase peptide resin **101** (Scheme 38) using HBTU.[67] Following this route would leave the less reactive amine (*para* conjugated with the amide) exposed to react to form the half receptor. Deprotection then reveals the 'more reactive' amine for the more challenging macrocyclisation step (Scheme 37).



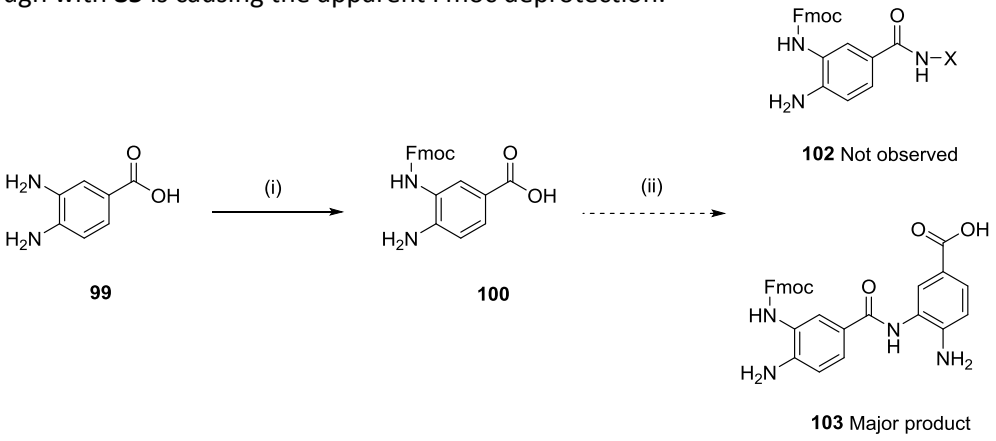
Scheme 37 Proposed synthetic route with protected diamino linker unit **95**. Where P = suitable protecting group, and R = solubilising group.



99 **100** **101**

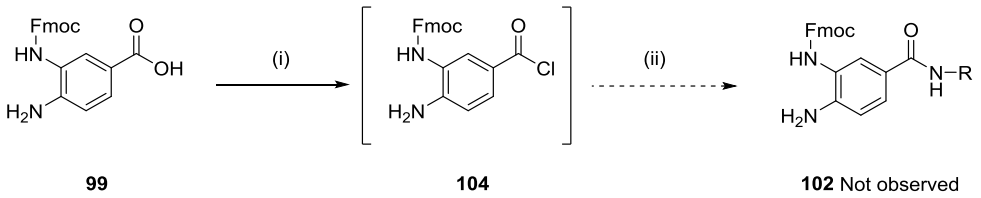
Scheme 38 Literature precedent for Fmoc-protection of one amine of 3,4-diaminobenzoic acid **99** and then subsequent amide coupling. R = solid phase peptide resin.

Fmoc protection of 3,4-diaminobenzoic acid **99** proceeded smoothly to give protected substrate **100** in excellent yield (Scheme 39). Initial attempts to replicate the amide coupling using G1M amine **85** as a test amine were unsuccessful. None of desired amide **102** was observed. Not all starting materials were consumed but the major product appeared to be dimerised product **103**, where apparent Fmoc deprotection of a percentage of starting material had caused the more reactive amine to be revealed and then participate in the coupling reaction to give a dimer. Other coupling reagents such as HATU and DCC were screened but gave either similar results or little evidence of reaction. The reaction was repeated using benzylamine as a test amine which afforded the benzyl amide in excellent yield. The reaction was then closely monitored by TLC and it was found that the undesired dimer **103** only formed upon addition of G1M amine **85**. This would suggest either the G1M amine **85** or an unknown impurity carried through with **85** is causing the apparent Fmoc deprotection.



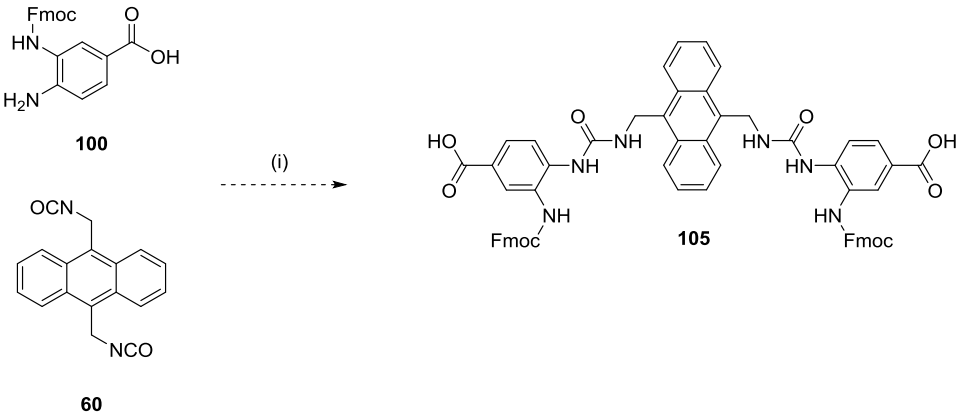
Scheme 39 (i) Fmoc-OSu, MeCN, $NaHCO_3$ (0.1 M), 16h, 90% (ii) G1M amine **85**, HBTU, HOBt, DIPEA, 16h, DMF.

Attempts to access the amide *via* the acyl chloride were also unsuccessful (Scheme 40). TLC analysis inferred that carboxylic acid **99** had been converted to the acyl chloride **104** with catalytic DMF in $SOCl_2$. Acyl chloride **104** was used in the next step without further purification. Addition of G1M amine **85** to **104** resulted in a complex mixture upon analysis by ^{1}H NMR of the crude reaction mixture, with none of the desired product isolated upon purification. Subjecting the acyl chloride **104** to the reaction conditions (no **85** present) and monitoring by TLC showed a second compound appearing over time, which could be a potential dimer as there is a free amine in the starting material. This would compete with the intended amide coupling reaction to yield a mixture of products.



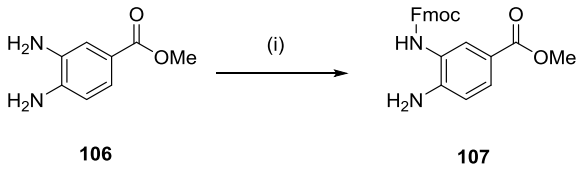
Scheme 40 (i) DMF, $SOCl_2$, (ii) G1M amine **64**, DMAP, Et_3N, CH_2Cl_2, rt, 16h.

With the direct coupling of G1M amine **85** to Fmoc-acid **100** proving to be an issue, a new route was attempted where the macrocycle or advanced intermediates were formed prior to attachment of the solubilising group. Addition of Fmoc protected acid **100** to isocyanate **60** was attempted and consumption of starting materials was observed by TLC, with a very fluorescent compound appearing on the baseline (Scheme 41). Purification on silica proved impractical due to poor solubility and polar functionality of molecules formed. Analysis by ^1H NMR of the crude reaction mixture suggested product formation (or at least reaction) had occurred with clear shifts of characteristic signals for both starting materials.



Scheme 41 (i) CH_2Cl_2, reflux, 3 days. **105** not isolated.

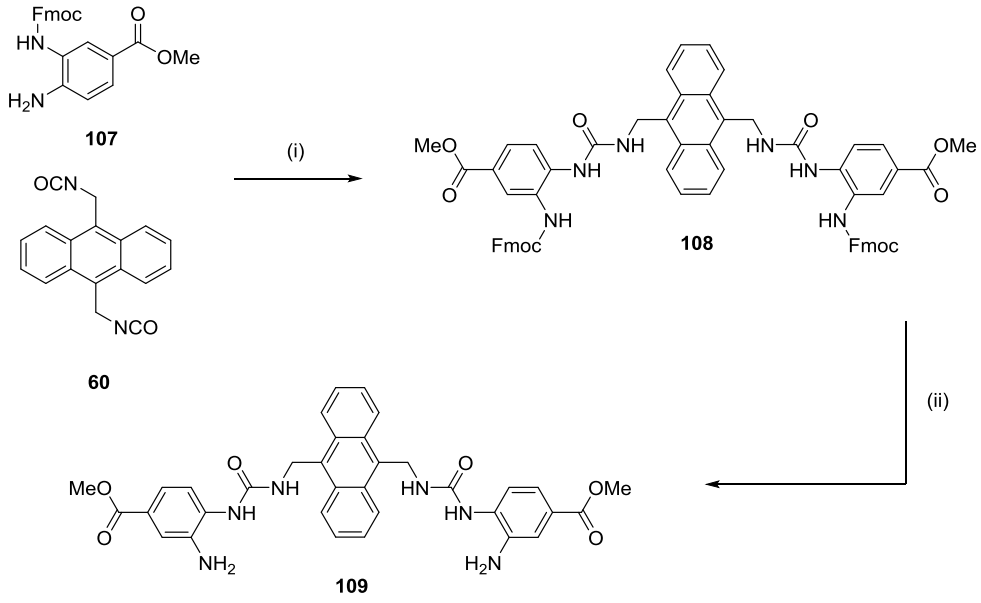
It was then investigated whether protection of the carboxylic acid groups would increase solubility and reduce the polarity of the molecule, thereby allowing purification by chromatography. The proposed route was then modified to use methyl 3,4-diaminobenzoate **106** instead of **100**, with the intention that the methyl ester would make the products more soluble and thus more manageable. Protection of methyl 3,4-diaminobenzoate **106** with Fmoc-OSu afforded the protected amino ester **107** in quantitative yield (Scheme 42).



106 **107**

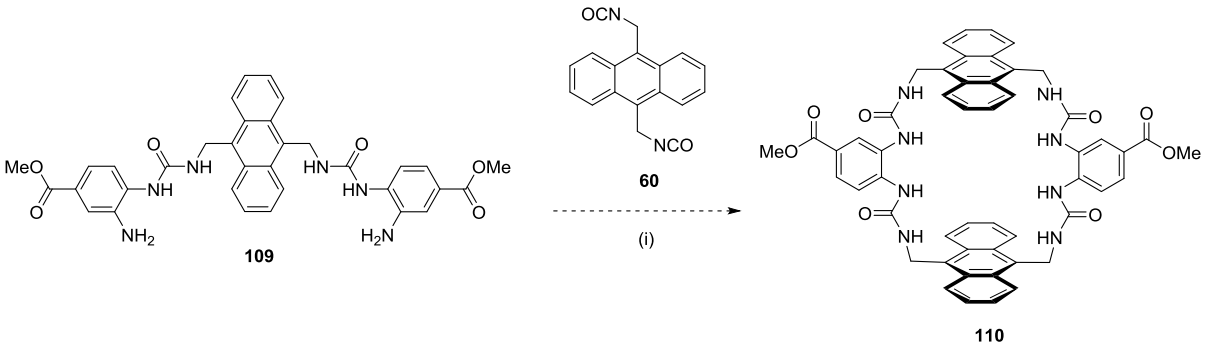
Scheme 42 (i) Fmoc-OSu, MeCN, $NaHCO_3$ (0.1M), 98%.

Subsequent reaction of **107** with isocyanate **60** afforded half receptor **108** in moderate yield (Scheme 43), although the product proved to be insoluble in most organic solvents (such as dichloromethane, methanol and ethyl acetate). Purification proved difficult due to insolubility in most eluents used for flash chromatography on silica and often caused the product to precipitate and adsorb to the silica stationary phase, and so **108** was used in the next step without further purification. Deprotection of half receptor **108** proved facile using piperidine in DMF (Scheme 43) and afforded amino half receptor **109** in excellent yield after trituration in dichloromethane.



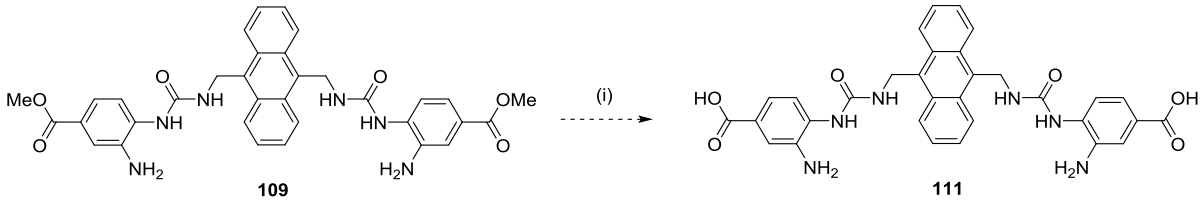
Scheme 43 (i) CH_2Cl_2, reflux, 3 days, 52%; (ii) piperidine, DMF, 1h, rt, 90%.

The macrocyclisation of half receptor **109** with isocyanate **60** was then attempted, using similar conditions that were successful for earlier system **37** (Scheme 44, Chapter 2). Consumption of isocyanate and half receptor was observed by TLC with new highly fluorescent compounds visible. However, any attempted purification proved very difficult due to very poor solubility of the crude reaction mixture. This also made analysis of the crude products difficult by ^{1}H NMR and reverse phase HPLC. Nevertheless, macrocycle **80** was observed by ESI mass spectrometry ($[M+2Na]^{2+}$ and $[M+3Na]^{3+}$ species), suggesting the macrocyclisation was successful to some degree.



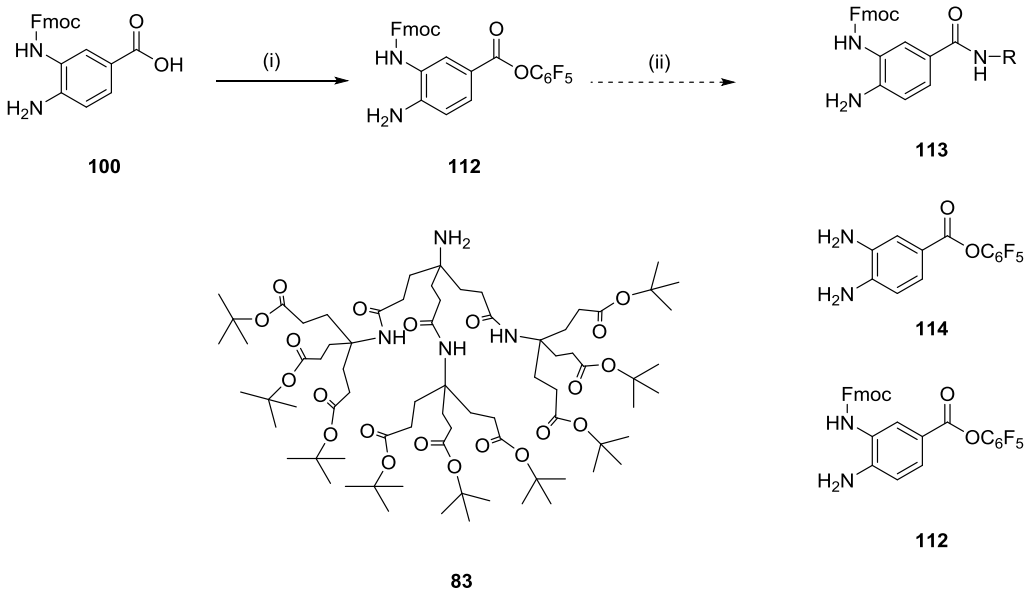
Scheme 44 (i) pyridine, CH_2Cl_2, rt, 4 days. **110** not isolated.

Attempts were then made to introduce the solubilising group at the half receptor stage, with the intention that this would aid purification of the macrocycle in the next step. Hydrolysis of the methyl ester to the carboxylic acid would be required to perform the amide coupling. Unfortunately, the very poor solubility of **109** proved sufficient enough to prevent the hydrolysis from proceeding, even when refluxed in a variety of solvent mixtures (THF and methanol), with >95% of starting material being recovered upon work up (Scheme 45).



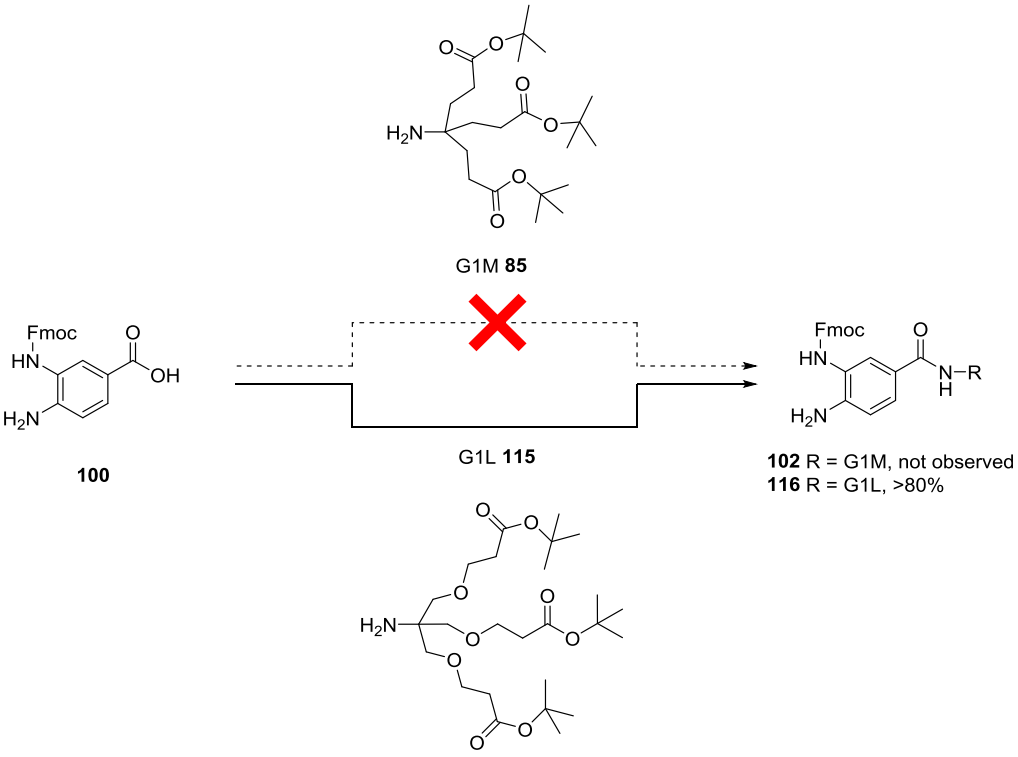
Scheme 45 (i) 2M NaOH (aq), MeOH, THF, reflux, 2 days. **111** not isolated.

Direct coupling of the solubilising group to the linker unit was then revisited, with Fmoc protected acid **100** converted to the pentafluorophenyl ester **112** using DCC and pentaflurophenol in good yield (Scheme 46). Ester **112** however proved remarkably unreactive, showing no reaction with G2MM amine **83** even after 2 weeks at reflux. The only observed change was a small amount of Fmoc deprotection to diamine **114**. This apparent lack of reactivity compared to previous pentafluorophenyl ester systems is most likely due to the more electron rich aromatic centre, as the free amine can delocalise its lone pair into the *para* located ester group. This would make the ester less electrophilic and thus potentially unreactive to the G2MM amine **83**, which may already be slow to react due to its large steric bulk.



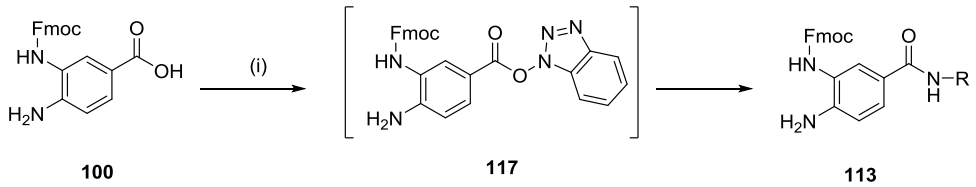
Scheme 46 (i) C_6F_5OH, DCC, THF, 84%, (ii) G2MM amine **83**, DIPEA, THF, reflux, 2 weeks, **113** not observed, **114** 5%, **112** >90%. R = G2MM solubilising group.

The direct amide coupling reaction with G2MM amine **83** and Fmoc acid **100** was then reconsidered, since it was becoming increasingly apparent during other studies in the Davis group that the G1M amine **85** was proving unreactive in known methodology. This methodology was very reproducible with several (including much larger) dendritic amines. Future work performed by Masters Student Kristian Jones would also demonstrate this unexpected lack of reactivity for G1M amine **85**. Several reaction conditions and coupling reagents were screened but the equivalent amide **102** was never isolated using G1M **85**. Remarkably, switching to the slightly larger "1st generation large" (G1L) amine **115** allowed access to the equivalent amide **116** on the first attempt in good yield. This would suggest that there is a fundamental problem with the nucleophilicity of G1M **85**; it could exist in solution in a self-associated conformation that masks the amino group and the dendrimer chain is the perfect size to facilitate this. This low reactivity of **85** could be the reason why most of the methodology described in this chapter was unsuccessful, as G1M **85** was used as a test compound given its more accessible synthesis compared to G2MM **83**. There was not time to repeat the aforementioned reactions with G2MM **83** however.



Scheme 47 Amide coupling reactions performed by Masters Student Kristian Jones, which revealed the fundamental lack of reactivity of G1M **85** towards forming aromatic amides. Similar dendrimer G1L **115** showed no indication of any such issues.

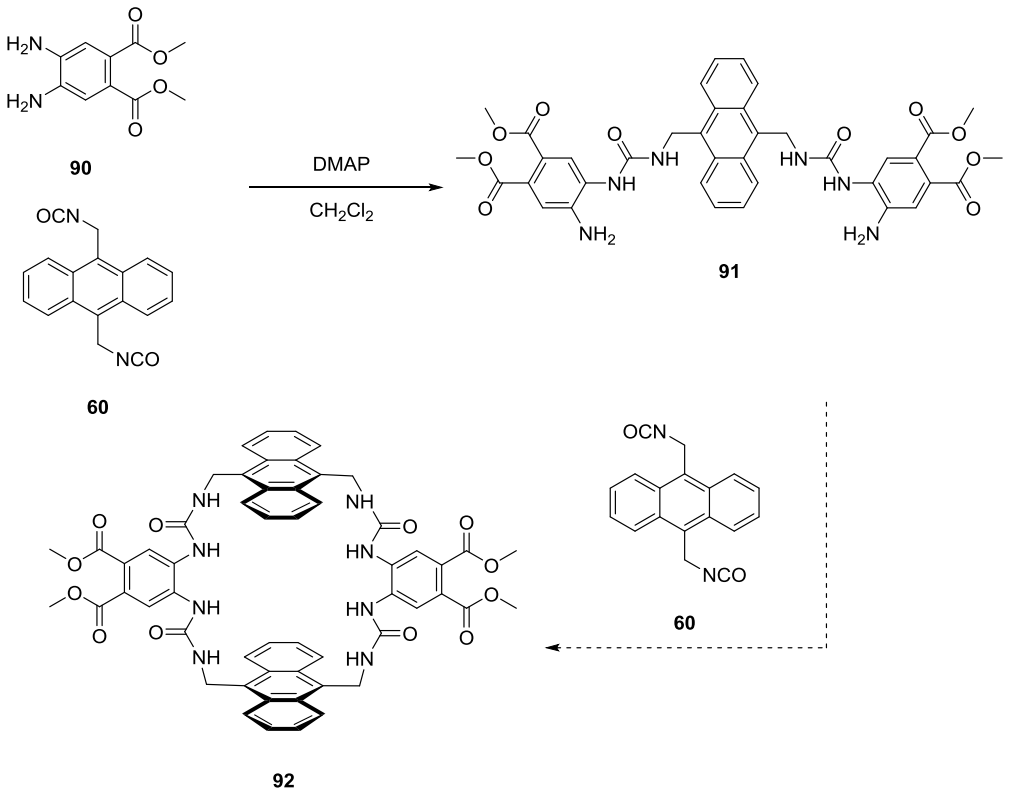
The amide coupling was then repeated using G2MM amine **83** and the conditions outlined in Scheme 39, with amide **113** isolated in good yield (76%) after purification by flash column chromatography. The reaction was further optimised by using THF in place of DMF as solvent, with a longer reaction time of 16 hours to achieve an improved yield of 86%. A precipitate was noticed during removal of the reaction solvent and addition of dichloromethane. Isolation and analysis of this precipitate identified it as benzotriazole ester **117**. It was then discovered that subjecting **100** to the reaction conditions without G2MM amine **83** formed this intermediate **117** exclusively, which could then be isolated by precipitation in a diethyl ether/water mixture and filtration. Ester **117** proved to be stable in the solid state to air and subsequent reaction with G2MM amine **83** afforded amide **113**. Test reactions indicated that **117** does undergo rapid hydrolysis when subjected to non-anhydrous reaction conditions however. There was no distinct advantage obtained in yield or purity of **113** by prior isolation of **117**, but it could prove a useful intermediate in future work if several linker derivatives are necessary.



Scheme 48 (i) G2MM amine **83**, HBTU, HOBt, DIPEA, THF, rt, 16h, 86%. R = G2MM solubilising group.

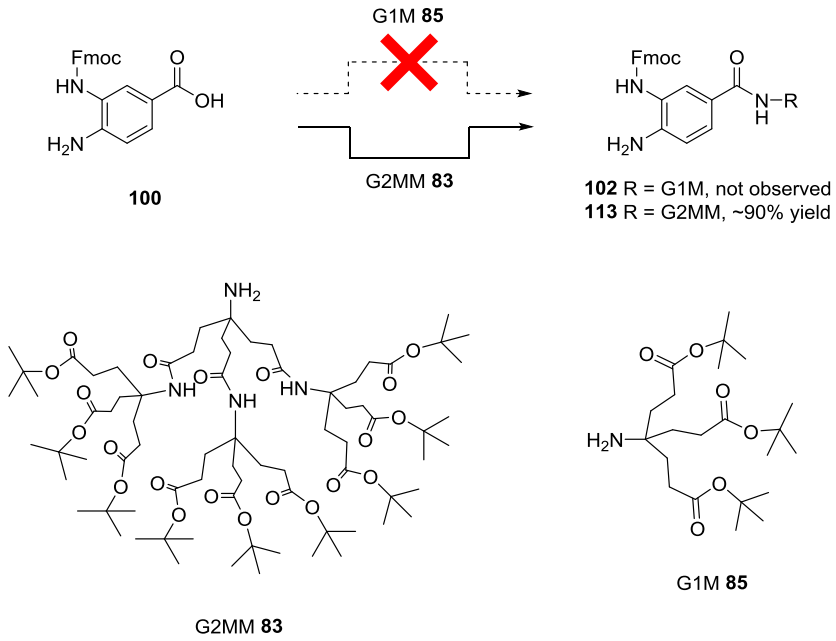
3.4 Conclusions

A synthetic route was successfully established towards diamino phthalimide **89**, which features large dendritic solubilising group G2MM. Initial attempts to react **89** with anthracene isocyanate **60** were unsuccessful and this synthetic route to the macrocycle was abandoned. Future work (Chapter 5) inspired a revisit to this reaction procedure however and it was found that addition of DMAP allowed the reaction of diamine **90** and isocyanate **60** to give half receptor **91** (Scheme 49). Unfortunately, there was not time to continue the synthetic route, but synthesis of macrocycle **92** seems entirely plausible given the successful synthesis of other anthracene macrocycles (Chapters 2 and 4). It is then envisaged that derivatisation of **92** will yield a library of receptors – including the originally targeted phthalimide receptor **81**.



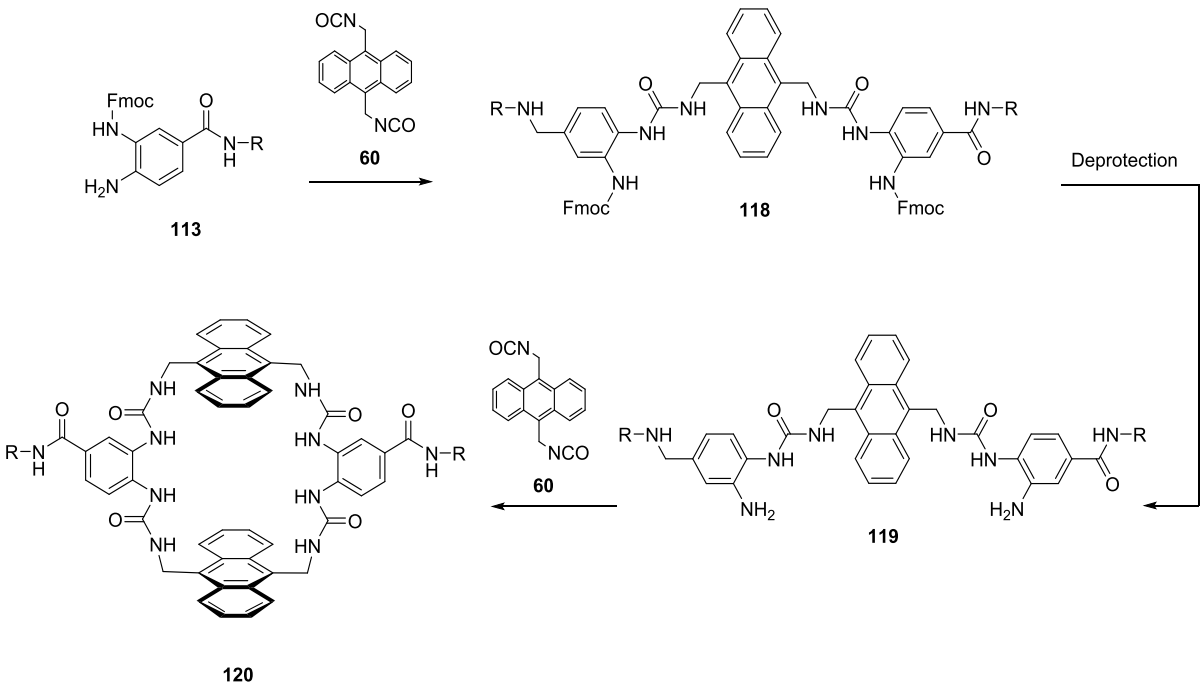
Scheme 49 Diamine **90** was synthesised successfully reacted with isocyanate **60**, in the presence of DMAP (inspired by future work, Chapter 5). There was not sufficient time to proceed this synthesis further however, but future of synthesis of macrocycle **92** is expected to be possible.

Efforts were then directed to simpler Fmoc protected linker **113**, in which the solubilising group was attached *via* a single amide. All test reactions of various methodology using G1M amine **85** were unsuccessful. It was eventually discovered that **85** appeared to have inherent issues with reactivity in amide coupling reactions with substrate **100** – as other dendritic and model amine substrates gave the equivalent amides in good yields. Fmoc-acid **100** was then converted to amide **113** upon reaction with larger amine G2MM **83**.



Scheme 50 All attempts to synthesise Fmoc-protected linker with G1M **85** were unsuccessful. Reaction occurred smoothly with G2MM **83** solubilising group however, suggesting an inherent issue with the reactivity of G1M **85**.

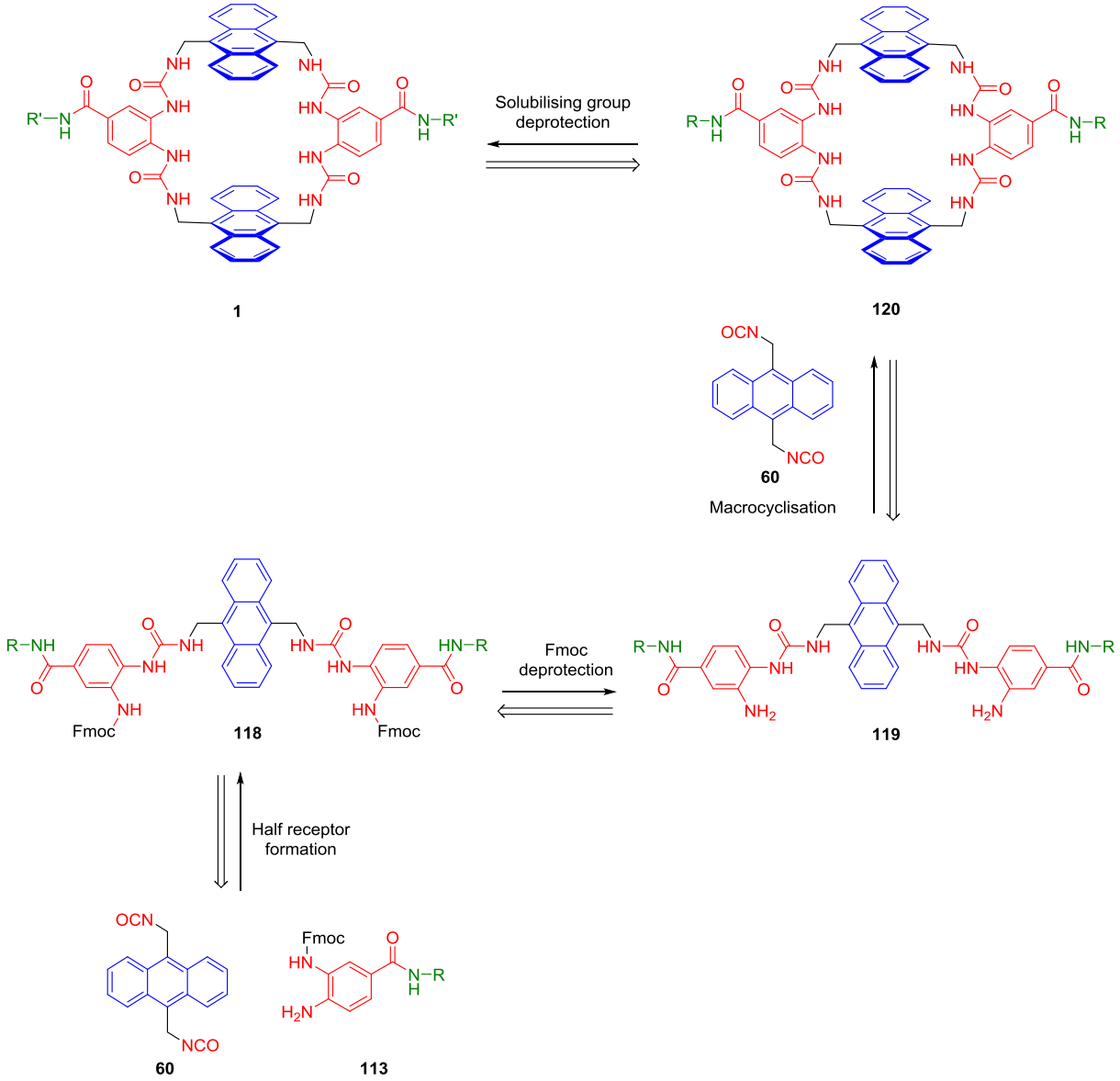
With target substrate **113** now in hand, the next step would be to form the half receptor **118** upon reaction with isocyanate **60** (Scheme 51). Fmoc deprotection then gives **119** followed by macrocyclisation to afford macrocycle **120**. The results obtained in carrying this out are covered in the next chapter (Chapter 4). The newly synthesised linker **113** has the potential to be incorporated into other receptors with different aromatic surfaces, other than anthracene. These receptors might possess novel and interesting activities towards certain guests. Indeed, such a receptor was synthesised and is discussed in Chapter 5.



Scheme 51 Proposed synthetic route towards macrocycle **120**, using Fmoc protected linker **113** and anthracene diisocyanate **60**.

Chapter 4 – Anthracene tetraurea receptor

4.1 Receptor design and retrosynthetic analysis

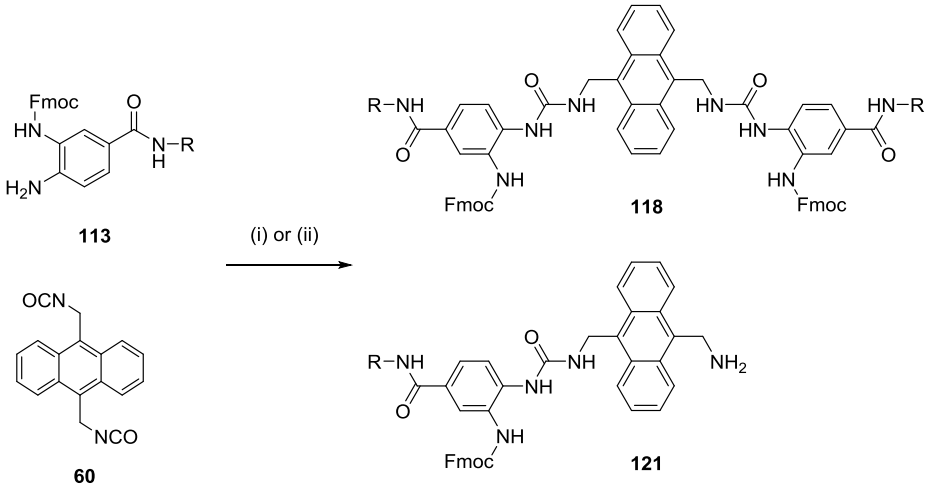


Scheme 52 Retrosynthetic analysis of anthracene tetraurea **1** leads to two key components, both of which were previously synthesised (chapters 2 and 3): isocyanate **60** and linker **113**. Formation of half receptor **119** and subsequent macrocyclisation yield protected macrocycle **120**. Global deprotection of the solubilising groups and neutralisation to pH 7 gives the water soluble receptor **1**. R = *tert*-butyl protected G2MM solubilising group. R' = water soluble G2MM solubilising group.

With a viable synthetic route established towards linker unit **113**, synthesis towards anthracene tetraurea **1** could proceed (Scheme 52). Formation of protected half receptor **118** followed by deprotection of the Fmoc groups affords amino half receptor **119**. Macrocyclisation by reaction of **119** with isocyanate **60** yields protected macrocycle **120**. Exposure of **120** to acidic conditions, followed by

neutralisation to pH 7, affords the desired polycarboxylate macrocycle **1**. It was hypothesised that use of a unsubstituted anthracene would circumvent the numerous instability issues encountered for previously synthesised anthracene macrocycles (Chapter 2) – most notably improved stability towards acid and oxidation by singlet oxygen in solution. Receptor **1** was designed towards binding linear cellodextrin oligosaccharides (reasons for which are covered in Chapter 2), and binding affinities for these substrates were to be determined using NMR spectroscopy titrations but also with fluorescence spectroscopy based methods, given the inherent fluorescent nature of anthracene.

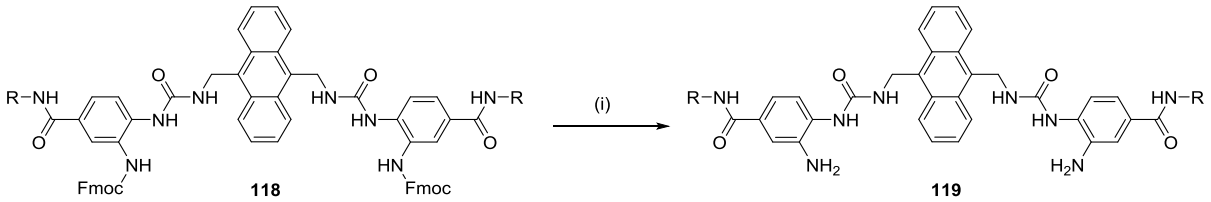
4.2 Synthesis and Characterisation



Scheme 53 (i) CH_2Cl_2, reflux, 2 days, **118** 67%, **121** ~20%; (ii) pyridine, CH_2Cl_2, reflux, 16h, **118** 86%, **121** not isolated. R = G2MM solubilising group.

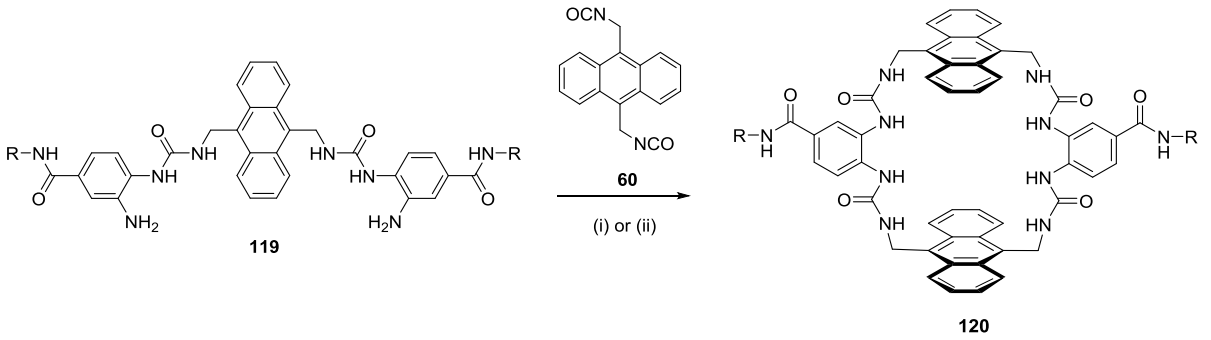
Synthesis of protected half receptor **118** was attempted by refluxing isocyanate **60** and linker **113** in dichloromethane. The reaction seemed slow, taking 2 days to appear to reach completion (Scheme 53). Separation of reaction products using silica as a stationary phase proved difficult, but separation was achieved using reverse phase HPLC with an acetone/water eluent, giving bis-urea **118** in good yield. Also isolated however was mono-urea **121** and linker **113** which appeared to be an indication of incomplete reaction. It was unclear whether **121** existed as the mono-isocyanate before purification by HPLC, as the water eluent could promote hydrolysis, or whether hydrolysis to the amine had already occurred during the reaction. The long reaction times suggested low reactivity for both intermediates though. Future work (chapter 5) discovered that addition of pyridine promoted a much faster and cleaner reaction when forming Fmoc protected half receptors from **113** and isocyanates. Application of this methodology to this system enabled an excellent yield of **118** with greatly increased rate of reaction, with no discernible amount of **121** produced either. Fmoc-protected half receptor **118** exhibited severe broadening of ^{1}H NMR signals during characterisation (in $CDCl_3$), most notably for signals relating to anthracene and linker aromatic protons. Signals related to the aromatic groups of

the Fmoc appeared relatively well resolved however. This broadening was observed in a variety of solvents, including more polar ones such as acetone and methanol, with dilution appearing to have no effect. It was therefore hypothesised that slow interconversion between conformations (relative to the NMR timescale) was occurring, potentially caused by slow rotation due to the bulky Fmoc groups. This severe broadening meant that full characterisation of **118** was not possible (especially ^{13}C characterisation), but the product was confirmed by high resolution ESI mass spectrometry.



Scheme 54 (i) DBU, CH_2Cl_2, 0 °C, 1h, 95%. R = G2MM solubilising group.

Deprotection of the Fmoc groups proved trivial using DBU in dichloromethane (Scheme 54), affording amino half receptor **119** in near quantitative yield after purification by column chromatography on silica. Half receptor **119** displayed severe broadening by ^1H NMR in $CDCl_3$, presumably due to intramolecular hydrogen bonding between ureas or intermolecular aggregation. Dissolution in CD_3OD appeared to disrupt these aggregates to give well resolved spectra however, allowing full characterisation.



Scheme 55 (i) CH_2Cl_2, reflux, 4 days, 23%; (ii) DMAP, CH_2Cl_2, reflux, 16h, 38%.

Reaction of amino half receptor **119** with anthracene isocyanate **60** was intended to afford the protected macrocycle **120**, however initial reactions in dichloromethane at room temperature were very slow. These were performed at relatively high dilution (<0.5 mM reaction concentration) to favour the intramolecular cyclisation, with a slow addition of isocyanate **60** to half receptor **119**. No reaction was observed after two days by TLC and HPLC however. Heating the reaction to reflux did appear to promote consumption of both starting materials however, and it was found that after 4 days all half receptor was consumed. Monitoring of the reaction by TLC was difficult due to the fluorescent nature

of anthracene, and so the reaction was monitored by reverse phase analytical HPLC. The presence of the target macrocycle in the crude reaction mixture was confirmed by mass spectrometry, which was deemed promising. Isolation of the macrocycle proved non-trivial initially however, with no separation achieved using flash chromatography on silica. Some separation was achieved using reverse phase preparative HPLC with an acetone/water eluent, but was hampered by consistent co-elution of unknown by-products. It was eventually found that initial purification by HPLC with an acetone/water eluent and then re-purification with a methanol/water eluent gave a very pure product (Figure 29). The closely co-eluting by-products were found to be oxidised forms of the receptor, where singlet oxygen had irreversibly reacted with the anthracene surfaces (see chapter 2 for a more detailed discussion on the oxidation of anthracene). This was possibly caused by the very long reaction times.

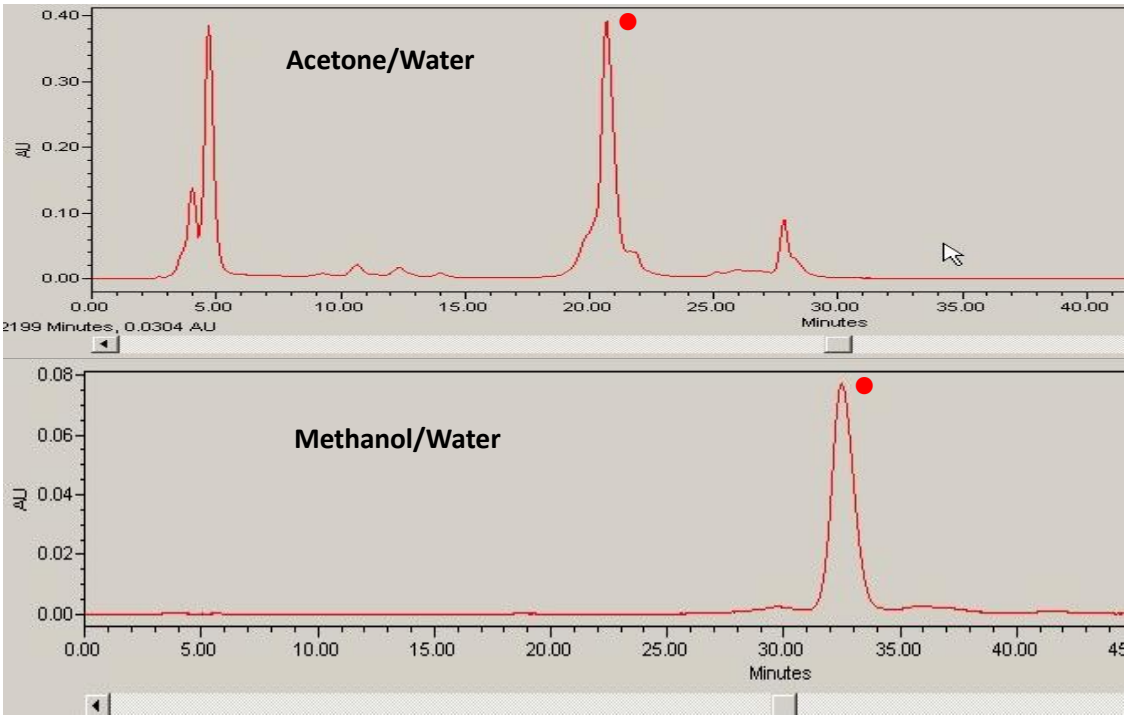


Figure 29 Reverse phase HPLC chromatograms for purification of macrocycle **120**. Separation of the product peak (denoted by •) and other impurities shown in acetone/water (top) and then methanol/water (bottom). Detection wavelength used λ = 394 nm. Stationary phase: preparative scale C18, flow rate: 16 mL/min, eluents: acetone/water (top) and methanol/water (bottom).

Future work (chapter 5) found that macrocyclisation rates were greatly enhanced by the addition of DMAP, and indeed the addition of DMAP appeared to have a very positive effect towards the formation of macrocycle **120**. Reaction times were vastly reduced to 16 hours, with an increased yield of ~40%. Purification was also markedly easier, with purification by reverse phase HPLC using only an acetone/water eluent to yield a very pure product. Macrocycle **120** was then fully characterised by ^{1}H and ^{13}C NMR, as well as high resolution ESI mass spectrometry, where the $[M+3Na]^{3+}$ and $[M+4Na]^{4+}$

adducts were observed. The effect that the asymmetry of the spacer unit has on the macrocycles symmetry was very apparent in the ^{1}H and ^{13}C NMR, showing two different signals relating to the anthracene surfaces (Figure 30).

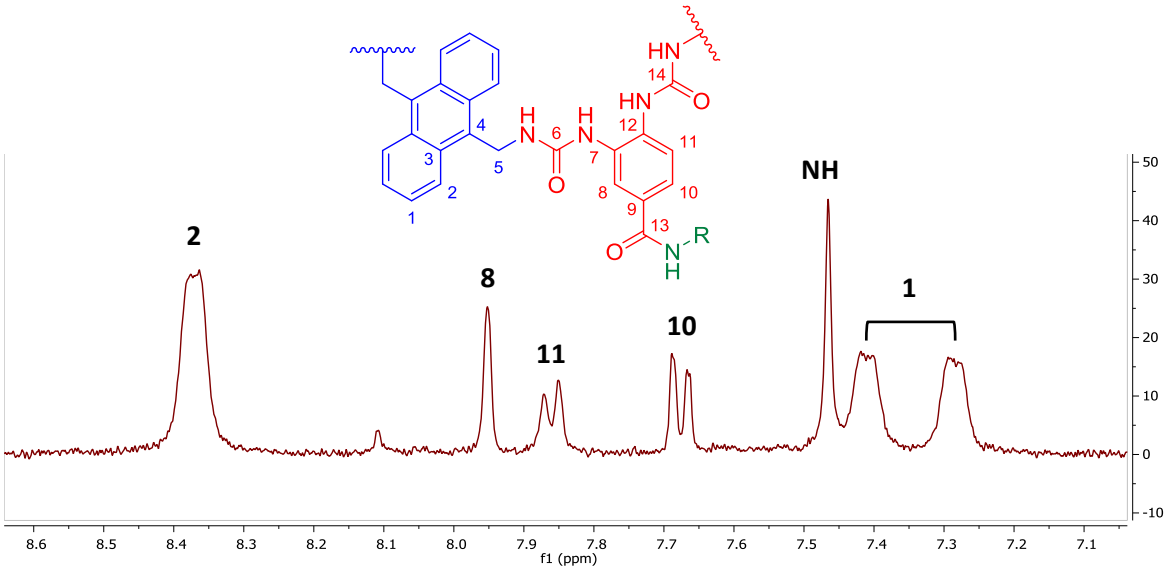
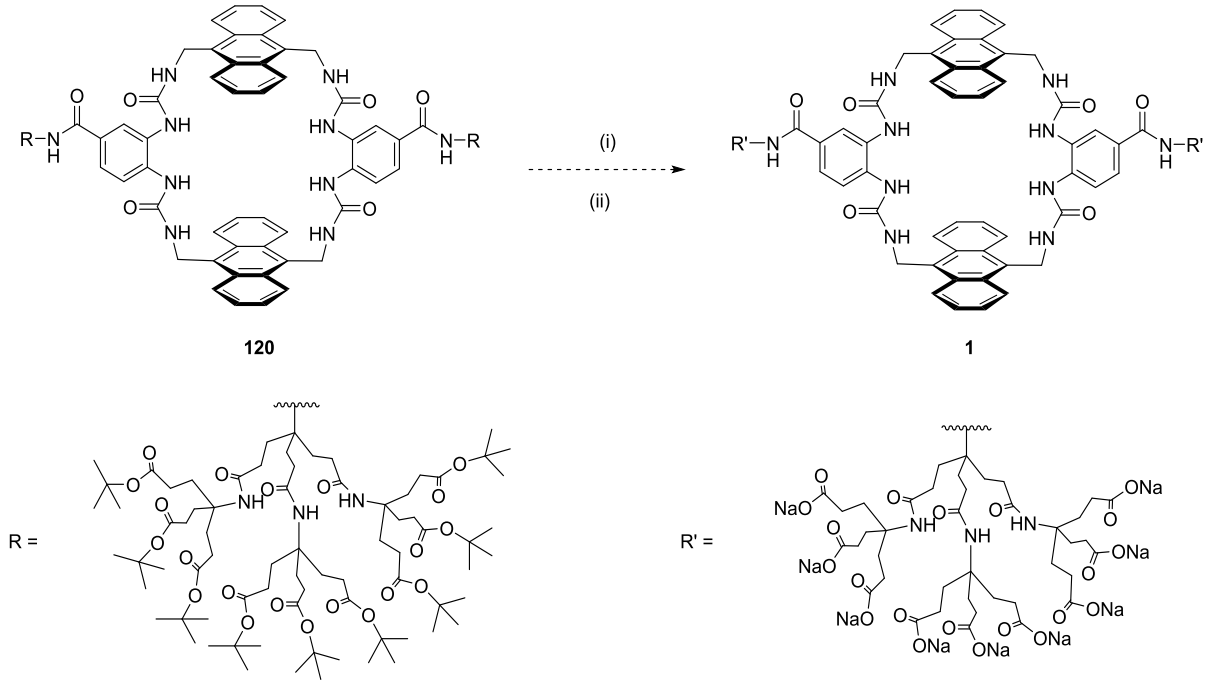


Figure 30 Partial ^{1}H NMR (500 MHz, CD_3OD, 298K) of macrocycle **120**, showing the aromatic region. The asymmetry of the spacer unit is apparent here, showing two different environments for anthracene protons at position 1.

As for previously discussed anthracene macrocycle **37** (chapter 2), tetraurea **120** showed susceptibility to oxidation with singlet oxygen in both solution and in the solid state. Solution stability was increased with the use of protic solvents such as CD_3OD, however storage of **120** under nitrogen in the solid state was necessary – as oxidation in the solid state in air was observed over a period of weeks, rendering the material impure and difficult to re-purify. It was therefore imperative to deprotect the solubilising groups to form the water soluble macrocycle **1**, where the aqueous solvent was expected to greatly suppress this oxidation.

Past anthracene receptors (with amide spacer units) that incorporated the G2MM solubilising groups used a minimum of 50% TFA (v/v) in dichloromethane for 16 hours to achieve complete conversion of the *tert*-butyl esters to the corresponding carboxylic acids. However, given the previous instability of urea anthracene macrocycles (discussed in chapter 2) to acid, such high concentrations of TFA were deemed too large a risk to the stability of macrocycle **120**. Thus, a more cautious approach was carried out.



Scheme 56 (i) TFA (20% v/v), CH_2Cl_2, rt, 16h; (ii) NaOH (aq). **1** not observed.

Unfortunately, exposure of macrocycle **120** to 20% TFA (v/v) in dichloromethane for 16 hours did not yield the corresponding polycarboxylic acid macrocycle. Removal of the TFA and solvent under vacuum afforded a crude solid that gave a suspension when water was added. This was to be expected as the polycarboxylic acid form of the receptor was predicted to be insoluble in water at low pH (i.e. pH <5), and was expected to become soluble upon neutralisation to pH 7 where to polycarboxylate form of the receptor **1** predominates. However, neutralisation of the crude mixture to pH 7 still yielded a suspension. Filtration of this suspension and analysis of the aqueous filtrate by ^{1}H NMR gave a spectrum that did not feature any signals relating to the anthracene surfaces (Figure 31). The only peaks observed were those characteristic of the spacer aromatic protons and the G2MM dendrimer solubilising group. Washing of the filtered solid with acetone and methanol yielded a fluorescent solution, analysis of which by ^{1}H NMR gave a spectrum consisting of two broad signals in the aromatic region. No signals pertaining to the spacer unit or dendrimer solubilising groups were observed, which would account for the poor solubility in water.

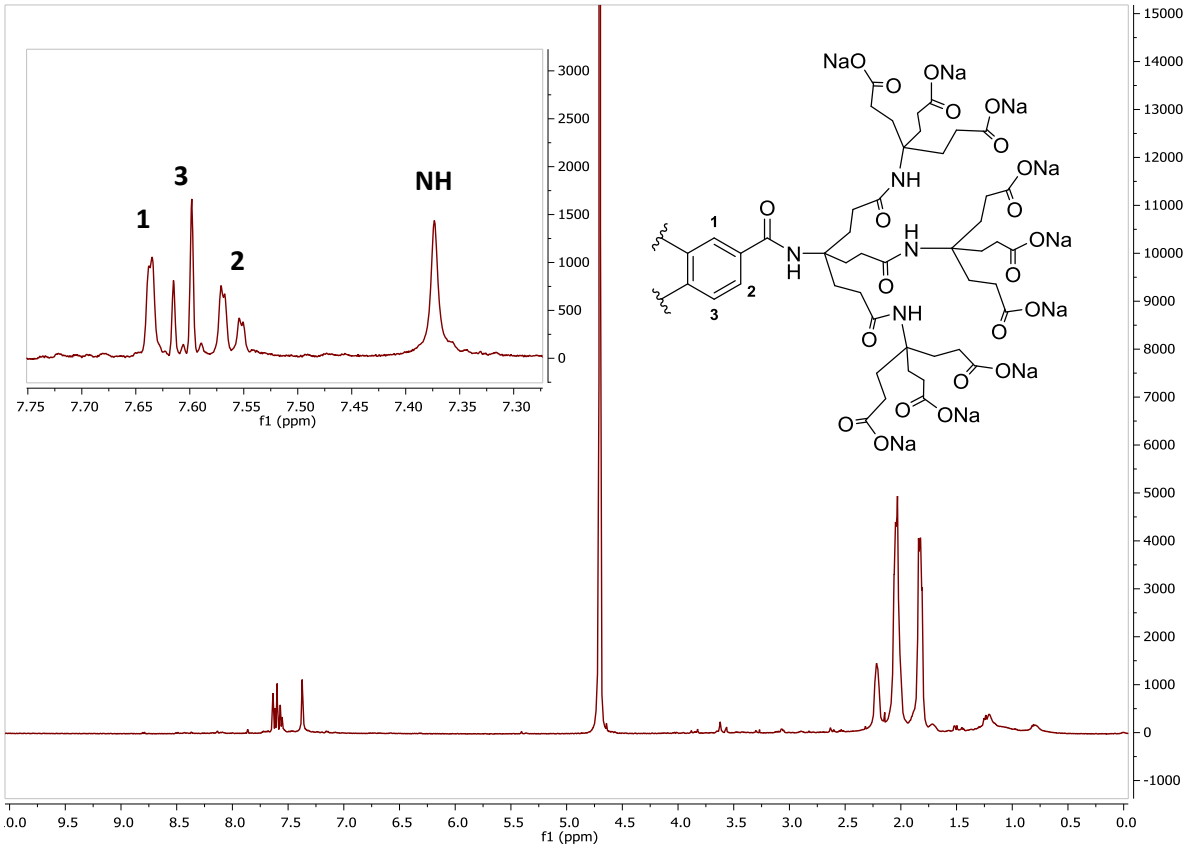
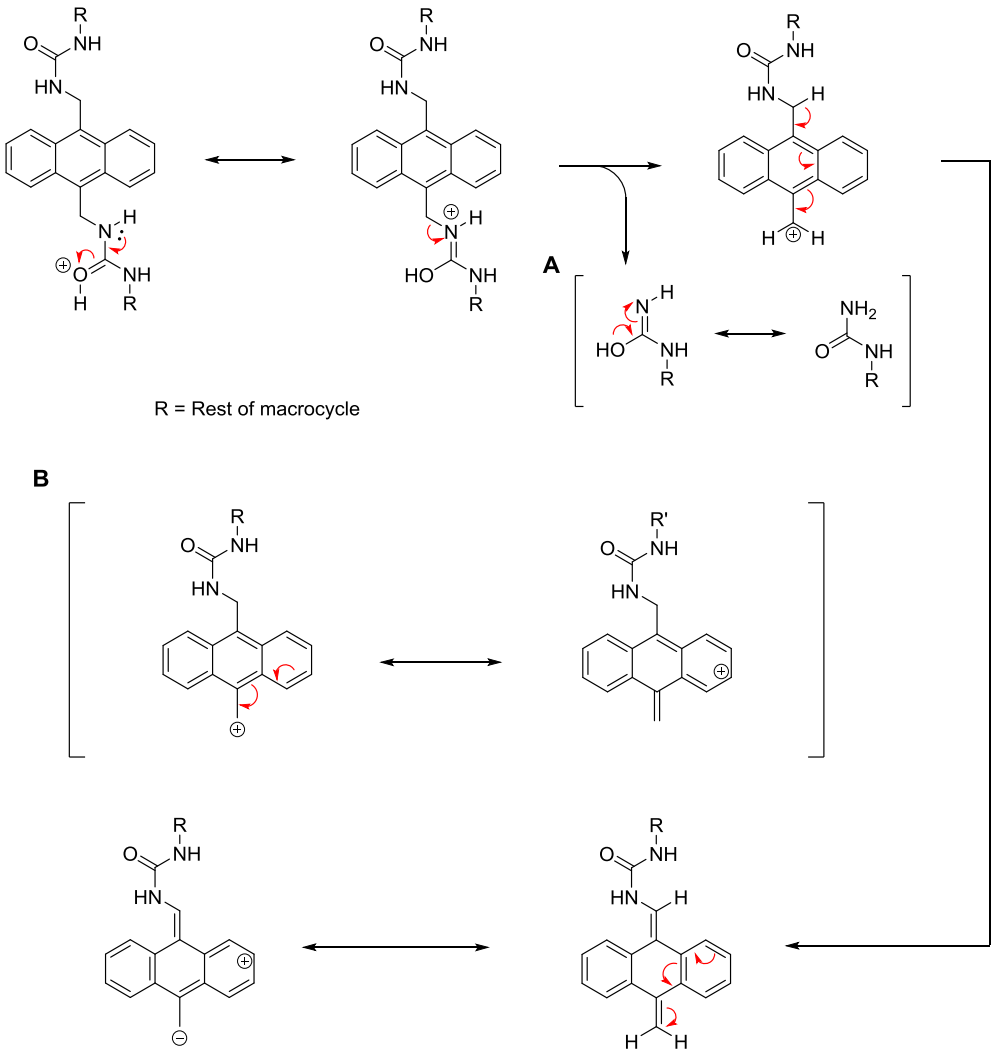


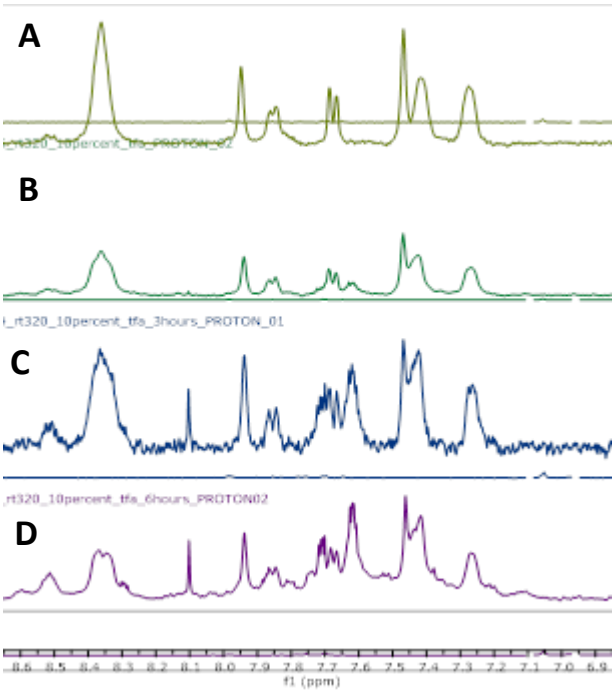
Figure 31 ^{1}H NMR (400 MHz, D_2O, pH 7, 298 K) spectrum of the water soluble fraction from the deprotection of macrocycle **120** with conditions outlined in Scheme 56. No signals relating to anthracene are observed, with only the characteristic signals for the spacer aromatic and dendrimer solubilising groups present. This would suggest that cleavage of the anthracene surfaces from the spacer units during the acidic deprotection was occurring.

These results would corroborate with the hypothesis discussed in chapter 2 about anthracene tetraurea macrocycle stability to acid, and suggest that an acid promoted heterolytic cleavage of the benzylic ureas was occurring – resulting in cleavage of the anthracene surfaces from the spacer units *via* a benzylic cation intermediate (Scheme 57). It had been speculated that use of the simpler anthracene (compared to the more electron rich derivatives described in chapter 2) would not favour this reaction pathway *via* the benzylic cation, but the initial evidence suggested this was not the case. However, it was presumed that the simpler anthracene might at least promote a slower decomposition of the macrocycle, with the loss of electron donating groups on the anthracene disfavouring formation of the benzylic cation. With this hypothesis in mind, various deprotection conditions were screened (mostly varying TFA concentration and reaction time), with the intention of discovering conditions that would drive full deprotection of the solubilising groups, but not decompose the macrocycle.



Scheme 57 Hypothesised mechanism for the acid promoted heterolytic cleavage of the benzylic ureas **(A)**, generating a resonance stabilised benzylic cation **(B)**. Loss of a proton generates a bis-alkene.

After screening a small range of conditions it was found that a minimum TFA concentration of 10% (v/v) was required to promote loss of the *tert*-butyls, with 5% TFA (v/v) showing no decrease in *tert*-butyl signals by ^{1}H NMR after 1 hour. Reaction of macrocycle **120** with 10% TFA (v/v) for 1 hour showed a 60% reduction in the *tert*-butyl signals by ^{1}H NMR (in CD_3OD) and no obvious signs of decomposition, which was deemed promising. However, increasing the reaction time to 3 hours still did not fully deprotect the solubilising groups. More concerning though was the appearance of several new signals in the aromatic and urea methylene regions of the spectrum, with peak shapes not resembling those related to the symmetrical macrocycle. Increasing the reaction time to 6 hours appeared to completely remove all *tert*-butyls from the solubilising groups, but also increased the intensity of the unknown signals.



A 5% TFA (v/v) for 1 hour.
No reaction observed.

B 10% TFA (v/v) for 1 hour.
~40% *tert*-butyls remain.

C 10% TFA (v/v) for 3 hours.
<5% *tert*-butyls remain.
New signals appearing in aromatic region

D 10% TFA (v/v) for 6 hours.
No *tert*-butyls remain.
Multiple species clearly present.

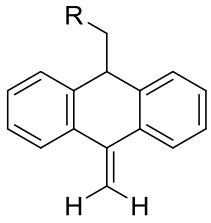
Figure 32 Partial ^{1}H NMR spectra (400 MHz, D_2O, pH 7, 298 K) of the crude reaction mixtures after reaction of protected macrocycle **120** with differing reaction conditions (listed right). Loss of the *tert*-butyl signal was observed over time but an increase of signals in the aromatic region also accompanied this (spectra shown left) suggesting some decomposition occurs before deprotection can reach completion.

Upon removal of the TFA and solvent, and neutralisation of the crude product to pH 7 with aqueous sodium hydroxide, a very fine suspension was observed – similar to the one seen when using the conditions described in Scheme 56. After filtration, analysis of the aqueous filtrate by ^{1}H NMR (in D_2O at pH 7) revealed a mixture of products had been obtained, potentially in similar ratios as no predominant signals for the symmetrical macrocycle were distinctly evident. Despite this, there were signals indicative of the symmetrical macrocycle, but there was no immediate procedure to separate the target macrocycle from these by-products at this time. What was apparent however, were potential signals relating to the hypothesised alkene by-products **122** at δ ~5.3 ppm in the ^{1}H NMR, which result from cleavage of the benzylic ureas to give a benzylic cation. Analysis of these suspected alkene signals using ^{1}H-^{13}C HSQC NMR revealed a ^{13}C signal of δ ~129 ppm, which is somewhat comparable to literature data for alkene compound **123** (Figure 33).[68] These data would compound the notion that acid promoted cleavage of the ureas was occurring.

Attempts to lower the acidity of TFA by changing the reaction solvent were also unsuccessful, with no reaction observed when using methanol as a solvent. Previous literature often uses >50% TFA (v/v) in methanol to convert *tert*-butyl esters to carboxylic acids, and this would explain the lack of reactivity observed. Use of weaker acids such as formic and acetic acid also showed no reaction, even at higher

concentrations of acid (>30% v/v). This would suggest that a stronger acid such as TFA was necessary to deprotect the *tert*-butyl esters, but seeing as this also resulted in decomposition of the protected macrocycle, it might not be possible to obtain conditions sufficient for both deprotection and macrocycle stability. It was therefore concluded from this study that even at these much milder deprotection conditions, rate of decomposition of the macrocycle was comparable to the rate of deprotection of the solubilising groups under acidic conditions, and that a clean outcome of this reaction was not possible. Given that no methods for purification of polycarboxylic acid compounds were also known at this time, it was decided to pursue a different protecting group strategy for the solubilising groups.

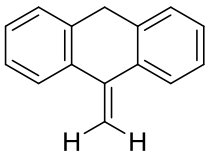
Hypothesised by-product



122

^{1}H δ: 5.32 ppm
^{13}C δ: 129 ppm

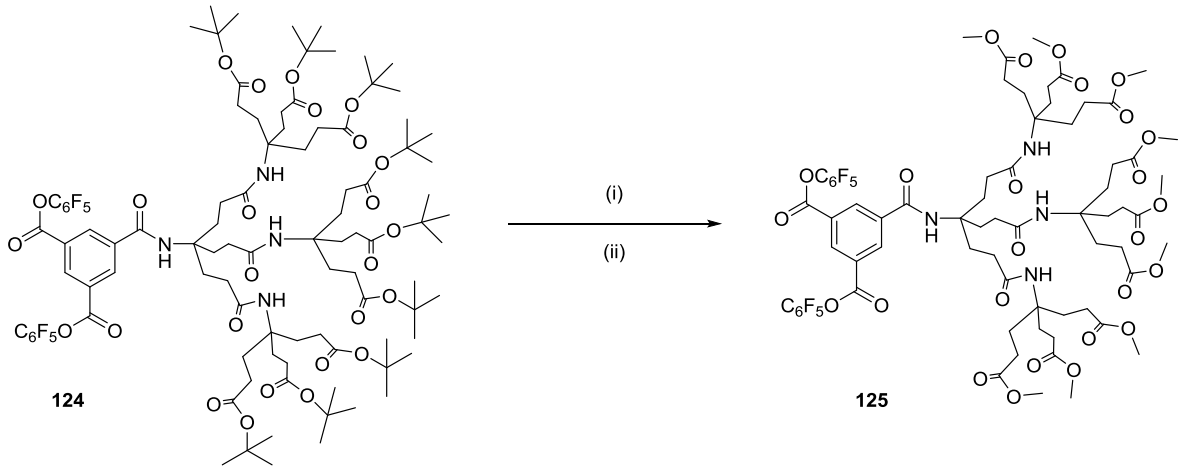
Literature compound



123

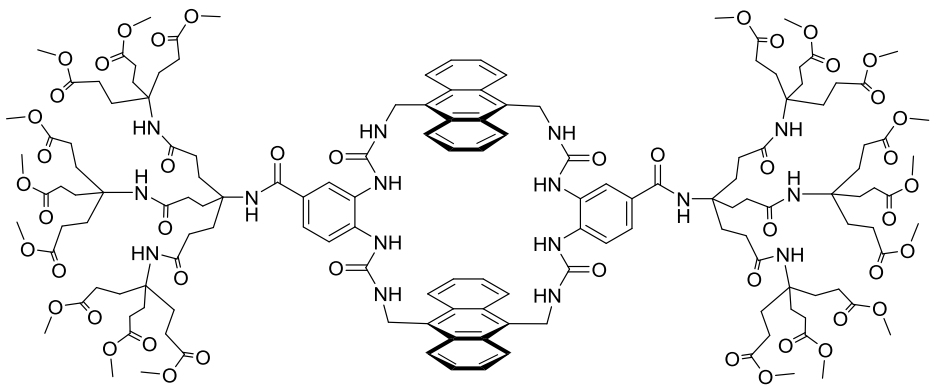
^{1}H δ: 5.55 ppm
^{13}C δ: 118 ppm

Figure 33 Hypothesised bis-alkene by-product **122** (left) and the observed chemical shifts (in D_2O) for the suspected alkene, which compares somewhat to a known literature compound **123** (right, chemical shifts quoted as for in $CDCl_3$).[68] This would suggest that the decomposition of macrocycle **120** under acidic conditions was proceeding *via* the pathway described in Scheme 57. R = rest of cleaved macrocycle.



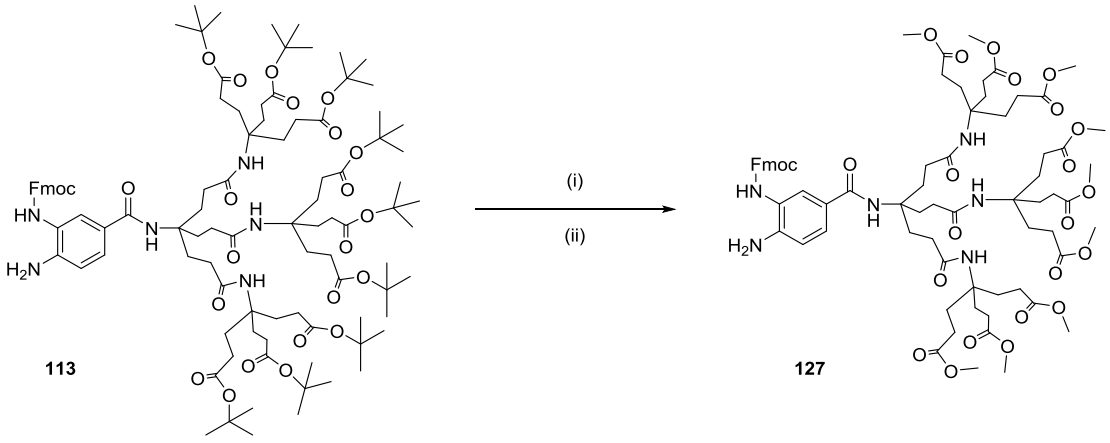
Scheme 58 (i) TFA, CH_2Cl_2, rt, 24h, (ii) 5% HCl (v/v), trimethyl orthoformate/MeOH (1:1), 24h, 99% over two steps.

After the investigation into the sensitivity of *tert*-butyl protected macrocycle **120** to acidic conditions, it was concluded that different carboxyl protecting groups be explored. Fortunately, previous work in the Davis group had developed methodology to convert linker **124** into the equivalent methyl ester version **125** in quantitative yield (Scheme 58).[39] This linker, once attached to a macrocycle, was then deprotected using very mild conditions: 0.5 M aqueous sodium hydroxide at room temperature. Published literature suggests that substituted ureas undergo hydrolysis very slowly unless subjected to very forcing conditions, and so it was deemed that the methyl ester protected macrocycle **126** would be a suitable synthetic target (Figure 34).[69] The conversion to the methyl ester proceeds via deprotection to the carboxylic acid using TFA and then esterification in acidic methanol with trimethyl orthoformate. This methodology was then applied to Fmoc protected linker **113** (Scheme 59), which was successfully converted to methyl ester linker **127** in excellent yield. The poorer yield when compared to **125** was attributed to potential reaction of the free amine in **127** reacting with methyl formate produced during the reaction.



126

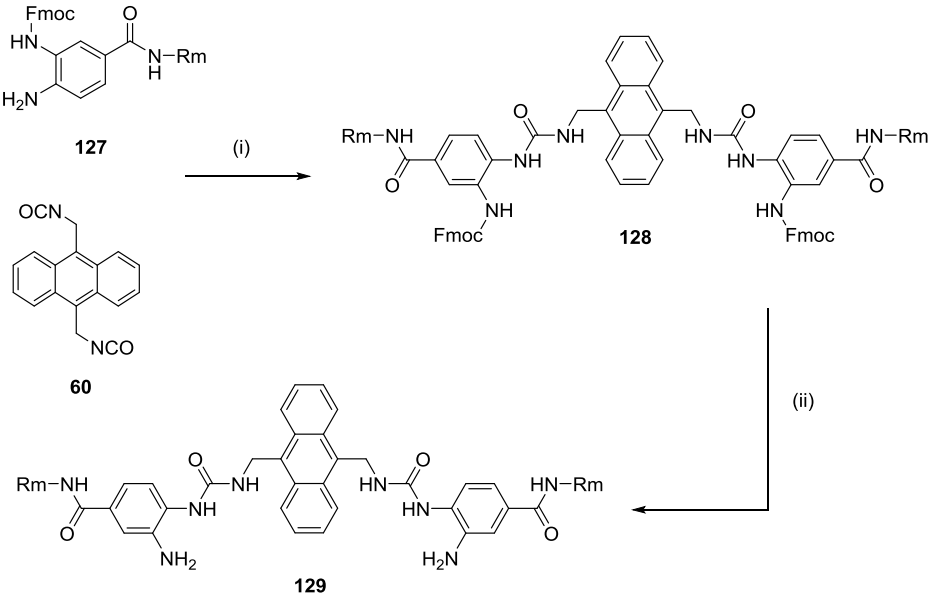
Figure 34 Structure of tetraurea anthracene macrocycle **126**, featuring methyl ester protected solubilising groups which afford the water soluble macrocycle **1** upon hydrolysis of the methyl esters with aqueous base.



113 **127**

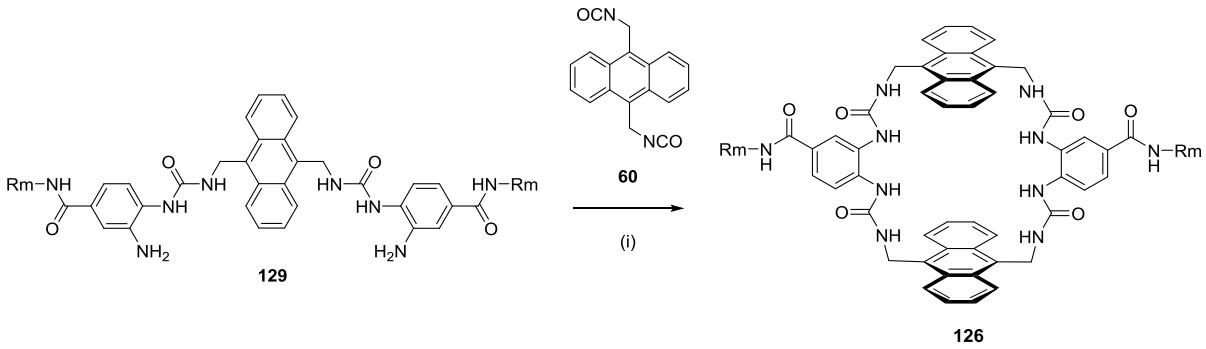
Scheme 59 (i) TFA, CH2Cl2, rt, 24h, (ii) 5% HCl (v/v), trimethyl orthoformate/MeOH (1:1), 24h, 87% over two steps.

Reaction of methyl ester linker **127** with isocyanate **60** proceeded as for previous half receptor **118** (Scheme 53), providing half receptor **128** in excellent yield (Scheme 60). An added benefit of the methyl ester protected dendrimers was that **128** proved easier to purify than equivalent *tert*-butyl half receptor **118** and separated well from excess linker **127** during purification by flash column chromatography on silica. This was presumably due to the increased polarity of **128** versus **118**. Deprotection of the Fmoc group to the diamino half receptor proved trivial again using DBU in dichloromethane, affording the product **129** in near quantitative yield (Scheme 60).



Scheme 60 (i) pyridine, CH_2Cl_2, reflux, 16h, 89%; (ii) DBU, CH_2Cl_2, 0 °C, 1h, 93%. Rm = methyl protected G2MM solubilising group.

The macrocyclisation of half receptor **129** and isocyanate **60** was then attempted, using the same conditions as for previous macrocycle **120** (Scheme 55, conditions ii). Upon completion, the reaction mixture was purified using reverse phase preparative HPLC with an acetone/water eluent. As for half receptors **128** and **129**, the methyl ester protected macrocycle **126** proved easier to purify than the equivalent *tert*-butyl protected macrocycle **120**, with better separation of peaks during HPLC purification. However, the yield was largely the same as for previous macrocycles, occurring in ~40% yield (Scheme 61). The main difference between *tert*-butyl **120** and methyl ester protected macrocycle **126** was solubility, where **126** appeared to only be soluble in DMSO. The ^{1}H and ^{13}C NMR (in $(CD_3)_2SO$) for **126** was very similar to that of **120**, and featured the same doubling of anthracene signals due to the asymmetry of the spacer unit.



Scheme 61 (i) DMAP, CH_2Cl_2, reflux, 16h, 38%. Rm = methyl protected G2MM solubilising group.

With methyl ester protected macrocycle **126** now in hand, deprotection of the solubilising groups with aqueous base was intended to yield the water-soluble macrocycle **1**. However, the very poor solubility of **126** proved a great hindrance when attempting hydrolysis of the methyl esters with aqueous sodium hydroxide, with no apparent conversion to the carboxylate when analysing the reaction mixture by ^{1}H NMR (in $(CD_3)_3SO$). Addition of polar organic solvents such as methanol or THF appeared to give improved solubility but the reaction still was unable to reach completion. Extensive analysis by analytical reverse phase HPLC was then undertaken to see whether separation of the full deprotected carboxylate could be achieved. Previous attempts in the group to purify polycarboxylic acid receptors with reverse phase HPLC were unsuccessful, and this was presumed to be due to the extreme polarity of the substrates giving no retention with the C18 stationary phase. However, it was found that acidifying the eluents with 0.1% TFA appeared to give retention. A gradient of 100% water to 100% organic solvent (usually methanol or acetonitrile) allowed separation and analysis of deprotection products. This method was then used to analyse the basic deprotection (in methanol and THF) of **126**, where it was very evident that the reaction had not reached completion – most likely due to the poor solubility of **126**. More than 14 distinct products were visible by HPLC, presumably due to varying degrees of methyl ester hydrolysis (Figure 35). The predominant products appeared to be the least polar as well, which would suggest that very little of the fully deprotected macrocycle had been synthesised.

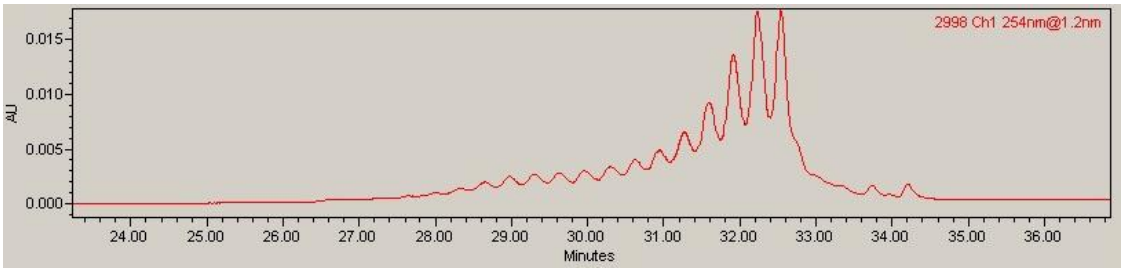


Figure 35 HPLC chromatogram of the reaction mixture after attempted hydrolysis of the methyl esters of macrocycle **126** with 0.5 M NaOH (aq) in methanol. A large distribution of products was evident, presumably due to varying degrees of methyl ester hydrolysis caused by poor solubility in the reaction medium. Detection wavelength of 254 nm used. Stationary phase: analytical scale C18, flow rate: 1 mL/min, eluents: methanol/water (+0.1% TFA).

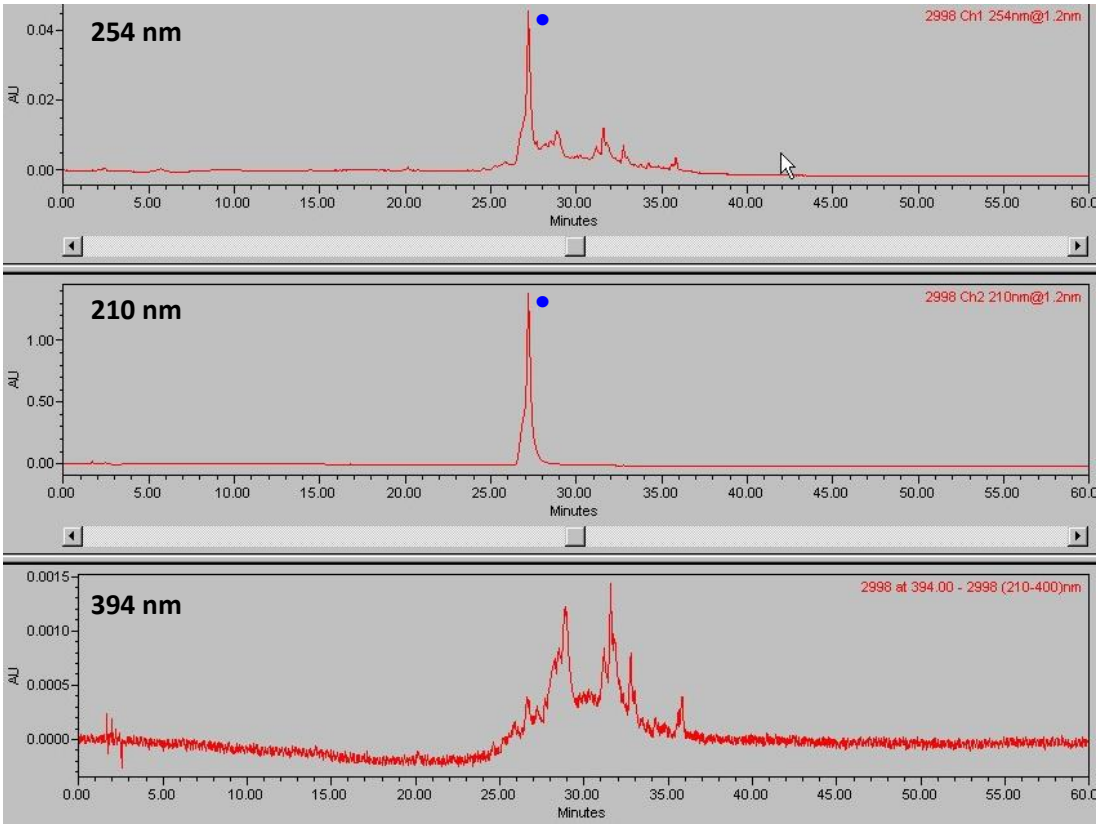
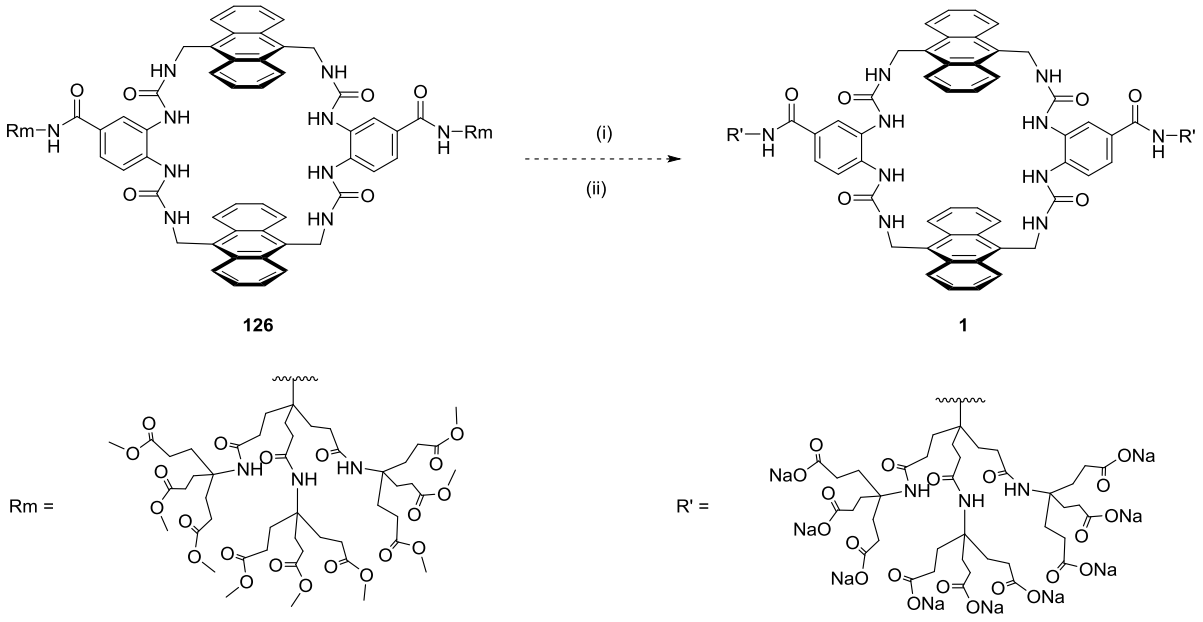


Figure 36 HPLC chromatograms of the reaction mixture after attempted hydrolysis of the methyl esters of macrocycle **126** with 0.5 M NaOH (aq) in DMSO. The predominant product (denoted with •) does not show UV absorption at 394 nm (the wavelength used to monitor for the presence of anthracene), which would imply this product does not feature anthracene. Analysis of this product by NMR confirmed this, as only signals relating to the spacer aromatic were observed, implying that hydrolysis of the ureas had occurred. A complex mixture of anthracene products was also observed (bottom). Stationary phase: preparative scale C18, flow rate: 16 mL/min, eluents: methanol/water (+0.1% TFA).

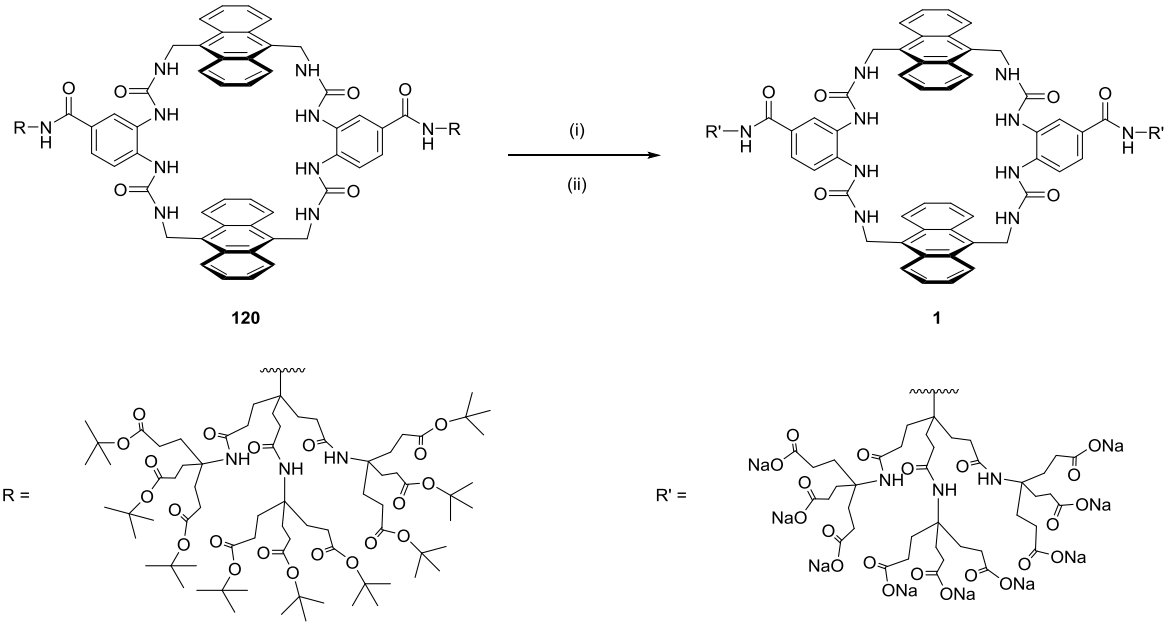
Given that improved solubility of macrocycle **126** was observed in more polar solvents such as DMSO and DMF, hydrolysis with 0.5 M sodium hydroxide in these solvents was attempted. While the improved solubility did appear to promote faster hydrolysis of the methyl esters, also apparently evident upon analysis of the reaction mixture by HPLC was hydrolysis of the ureas as well. The major product isolated from the HPLC did not shown UV absorption at 360-400 nm, which was the wavelength typically used to monitor for the presence of fluorescent compounds such as anthracene. Analysis of this product by ^{1}H NMR (in DMSO) inferred the product to be the spacer unit, with fully deprotected solubilising group, now no longer attached to anthracene – an indication full hydrolysis of all the ureas had occurred. Attempts to lower the concentration of sodium hydroxide used had little positive effect, either rendering no reaction or giving partial hydrolysis of the ureas to give a complex mixture of products by HPLC (Figure 36). Ultimately, very little control over the hydrolysis was achieved. This was unexpected given that published research suggested the rates of urea hydrolysis are very slow.[69] It was theorised that the rate of urea hydrolysis for **126** might be greater than for acyclic ureas, as this relieves macrocyclic strain.



Scheme 62 (i) 0.5M NaOH (aq), DMSO, 2h; (ii) Amberlyst 15 acidic ion exchange resin, H_2O, to pH 6 then NaOH (aq) to pH 7.4. Receptor **1** not isolated.

Conversely, more reaction control was exhibited when attempting the deprotection of the *tert*-butyl protected macrocycle **120** with TFA (Figure 32). However, there was no viable method for separation of the deprotected macrocycle at that time. Therefore, with methodology now developed that enables purification of polycarboxylic acid macromolecules by HPLC, the acidic deprotection of anthracene

macrocycle **120** was revisited with the intention of purifying the deprotected macrocycle after closely monitoring the progress of the reaction by analytical HPLC. Exposure of *tert*-butyl protected macrocycle **120** to 10% TFA (v/v) in dichloromethane for 2 hours, and then analysis of the crude reaction by analytical HPLC revealed 4 well separated major products. Isolation of these products by preparative reverse phase HPLC and analysis by ^{1}H NMR (CD_3OD) revealed most products still featured signals relating to the *tert*-butyl esters, implying the reaction had not reached completion. Increasing the reaction time to 4 hours appeared to generate one predominant product, eluting at ~28 minutes (Figure 37). Analysis of the earlier eluting products (~26 and 27 minutes) suggested this was a decomposition by-product, whereas later eluting products still featured *tert*-butyl esters and are likely partially deprotected macrocycles. The predominant product featured no NMR signals relating to decomposition or *tert*-butyl esters, and so was theorised to be the target polycarboxylic acid macrocycle, which was isolated in ~80% yield.



Scheme 63 (i) TFA (10% v/v), CH_2Cl_2, rt, 4h; (ii) NaOH (aq), H_2O, 80% over two steps.

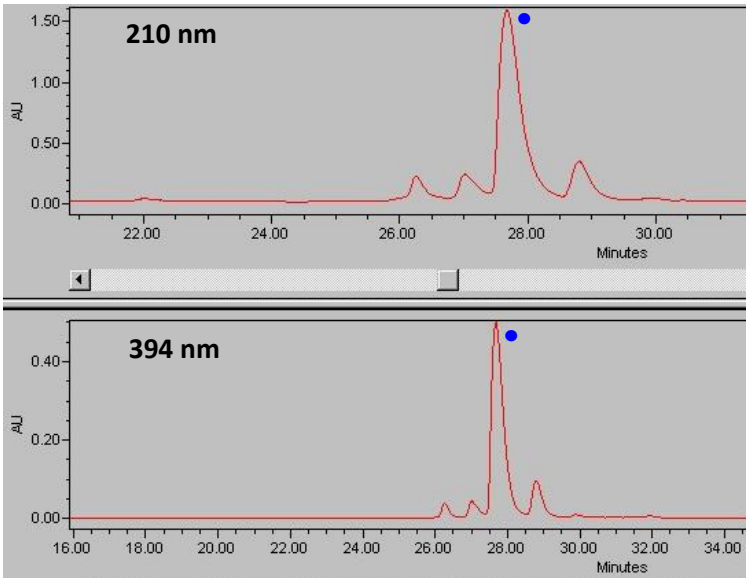


Figure 37 Reverse phase HPLC chromatograms for the purification of **120** using conditions outlined in Scheme 63, using a methanol/water (0.1% TFA) eluent. The major product at ~28 minutes (denoted with •) was desired deprotected macrocycle **1**, with earlier products resulting from acid promoted decomposition. Later eluting compounds are partially deprotected macrocycles. Receptor **1** isolated in 80% yield. Detection wavelengths used are 210 nm (top) and 394 nm (bottom). Stationary phase: preparative scale C18, flow rate: 16 mL/min, eluents: methanol/water (+0.1% TFA).

Neutralisation of the polycarboxylic acid macrocycle to pH 7.4 with aqueous sodium hydroxide was expected to yield the water soluble macrocycle **1**. Indeed, at pH 5 in water a suspension was observed, which became a colourless solution upon approaching pH 7. Freeze drying this solution and reconstitution in D_2O allowed analysis of tetraurea **1** by ^1H NMR. However, analysis of **1** in D_2O indicated that signals relating to the anthracene and urea methylenes had significantly broadened, almost to the point where they were no longer discernible (Figure 38). Only signals relating to two of the spacer aromatic protons and the dendrimer solubilising groups were well resolved. A dilution study was performed to ascertain whether this broadening was due to intermolecular aggregation, but spectra over the concentration regime 2-0.1 mM were consistent implying that this was not the case. It was then postulated that the broadening could be a result of slow rotation about bonds or interconversion between conformations. A variable temperature study by ^1H NMR was then performed, and indeed reaching 45 °C and above showed a dramatic sharpening of all ^1H NMR signals relating to the core macrocyclic structure (Figure 39). This would corroborate with the hypothesis that receptor **1** undergoes slow interconversion between conformations at room temperature, and the elevated temperatures increase this rate of interconversion to give more resolved spectra.

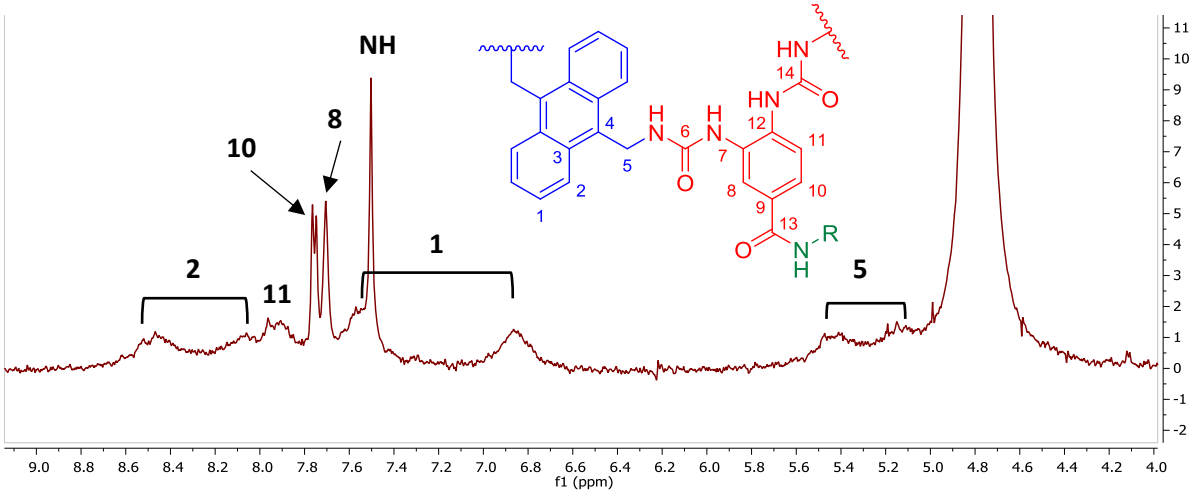


Figure 38 Partial ^{1}H NMR spectrum (500 MHz, D_2O, pH 7.4, 298 K) of receptor **1** (0.5 mM) in. Majority of signals relating to core macrocyclic structure are significantly broadened at 298 K, with protons 8 and 10 (and NH) being the only well resolved signals. Dilution of the receptor appeared to have no effect, which implied the broadening was not due to intermolecular aggregation.

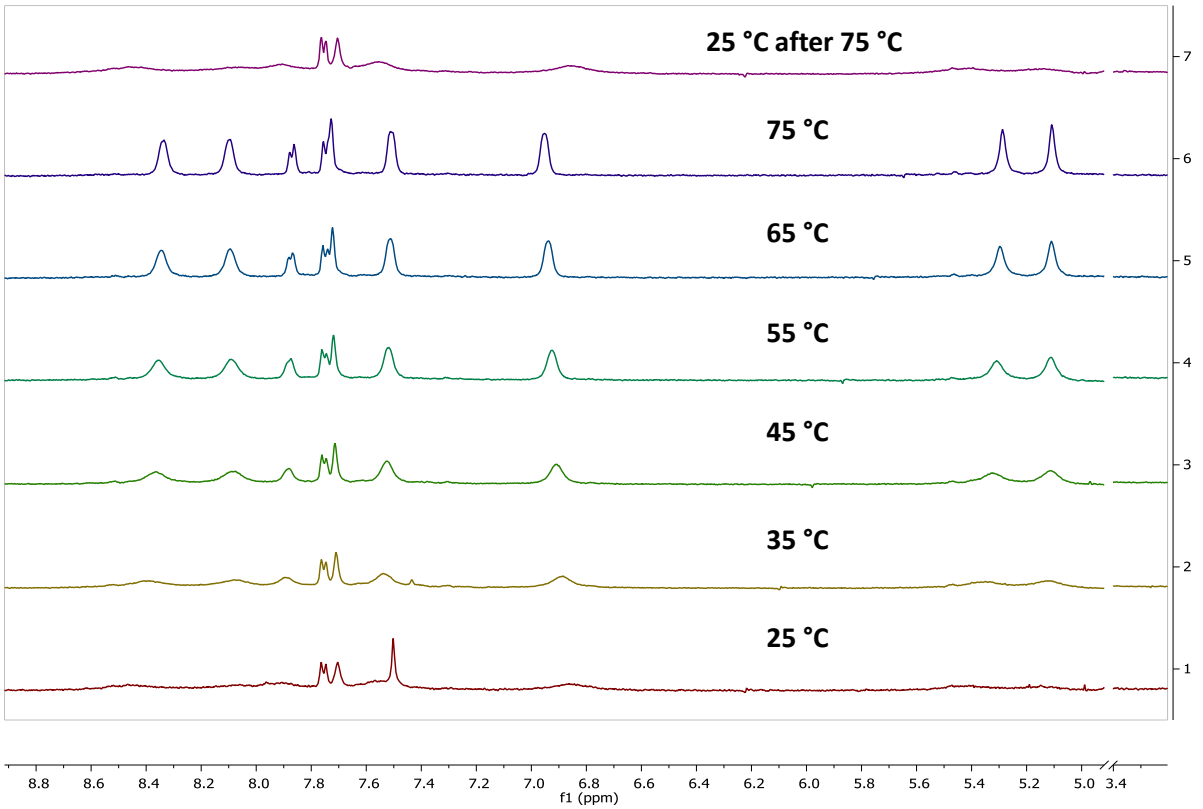
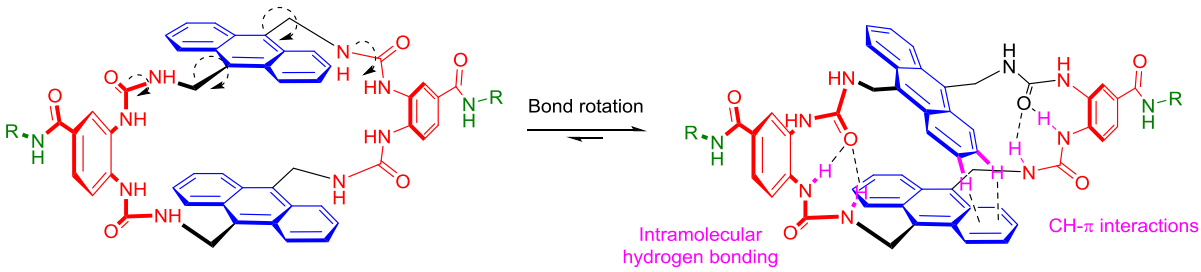


Figure 39 Partial ^{1}H NMR spectra showing receptor **1** at various temperatures in D_2O. Receptor signals sharpen upon heating – most notably for signals relating to anthracene and urea methylenes.

The origins of this slow interconversion between conformations could be due to intramolecular hydrogen bonding from the ureas on each spacer unit, as well as hydrophobic CH-π interactions between each anthracene surface which would be strengthened in the aqueous solvent (possible

conformation proposed in Scheme 64). Evidence for this can be seen in the ^{1}H NMR, where one of the anthracene protons at position 1 (δ ~6.8 ppm) was much more shielded and thus upfield than the others – due to the protons interacting with the anthracene π-cloud. The fact that four distinct protons environments for the protons at position 1 are not seen, suggest that the receptor is interconverting between conformations faster than the NMR timescale and what was observed is the time averaged NMR spectrum. Although the severe broadness of peaks might suggest that interconversion rate is similar to the NMR timescale (fast-intermediate exchange rate), and increasing the temperature increases this rate of interconversion to give rise to the more resolved signals observed (fast exchange rate).



Scheme 64 Proposed conformations of receptor **1** in water at 298 K, with the conformation on the right predominating. Bond rotation enables preference for the right-hand conformation through intramolecular hydrogen bonding between ureas and CH-π interactions between the anthracene surfaces.

This 'self-associated' conformation (Scheme 64) was also observed through molecular modelling of the ground state conformation of receptor **1** using MCMM calculations (Figure 40). Where it was predicted that the ground state conformation featured intramolecular hydrogen bonding between ureas and rotation of one anthracene surface to facilitate a CH-π interaction with the protons at position 1. Further work is currently underway to fully characterise the conformation of the free receptor **1** in aqueous media through in-depth NMR studies such as TOCSY and NOESY in a 9:1 H_2O/D_2O solvent composition, which would allow characterisation of interactions with the urea NH protons. NOE interactions between the urea NHs and spacer aromatics would infer that rotation of the ureas is occurring as described in Scheme 64. Differentiation of each anthracene surface would also allow any observed interactions to be absolutely characterised and allow construction of a more confident hypothesis regarding conformation. This could be achieved through TOCSY NMR to ascertain which NMR signals correspond to each anthracene surface, NOE interactions could then be assigned to each surface confidently and see if they correlate to the proposed conformation described in Scheme 64.

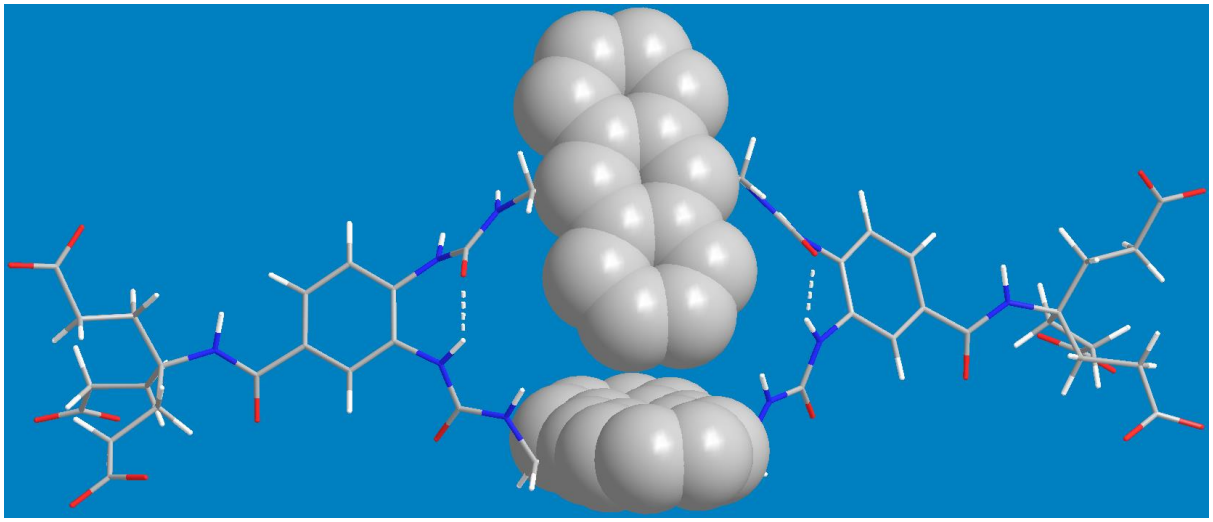


Figure 40 Ground state conformation of receptor **1** from Monto Carlo molecular mechanics conformational searches. Observed are intramolecular hydrogen bonds between ureas (dashed lines) and rotation of one anthracene surface to facilitate CH-π interactions. This corroborates with information from NMR that the predominant conformation of receptor **1** is that as described in Scheme 64. Force field used: OPLS2005 with aqueous GB/SA solvation.

4.3 Binding Studies

Despite indications during characterisation that receptor **1** potentially featured a collapsed cavity due to intramolecular interactions between ureas and anthracene surfaces, binding studies with several carbohydrate substrates were undertaken using a range of analytical techniques. If binding of a particular guest molecule is more favourable than the predicted intramolecular interactions, then binding events would be observable and thus association constants determinable. Initial studies were directed towards smaller substrates such as monosaccharide D-glucose **4**. This was to determine what differences in binding affinity the urea spacer unit enabled over previous anthracene receptor **25**, which utilised an isophthalamide spacer to give an association constant of K_a ~60 M^{-1} for D-glucose **4**.[35]

An NMR titration in D_2O was therefore undertaken, where both stock solutions of receptor **1** and substrate (here D-glucose **4**) were confirmed to be at pH 7.4 to avoid changes in pH giving misleading changes in the NMR spectrum upon addition of D-glucose **4**. The solution of D-glucose **4** was allowed to reach equilibrium of both α and β anomers by dissolution in the titration medium (here D_2O) 24 hours beforehand, with the equilibrated anomeric ratio verified by ^1H NMR. A solution of D-glucose **4** (1 M), with added receptor **1** (0.5 mM) to keep the concentration of receptor constant throughout the titration, was titrated into a solution of receptor **1** (0.5 mM).

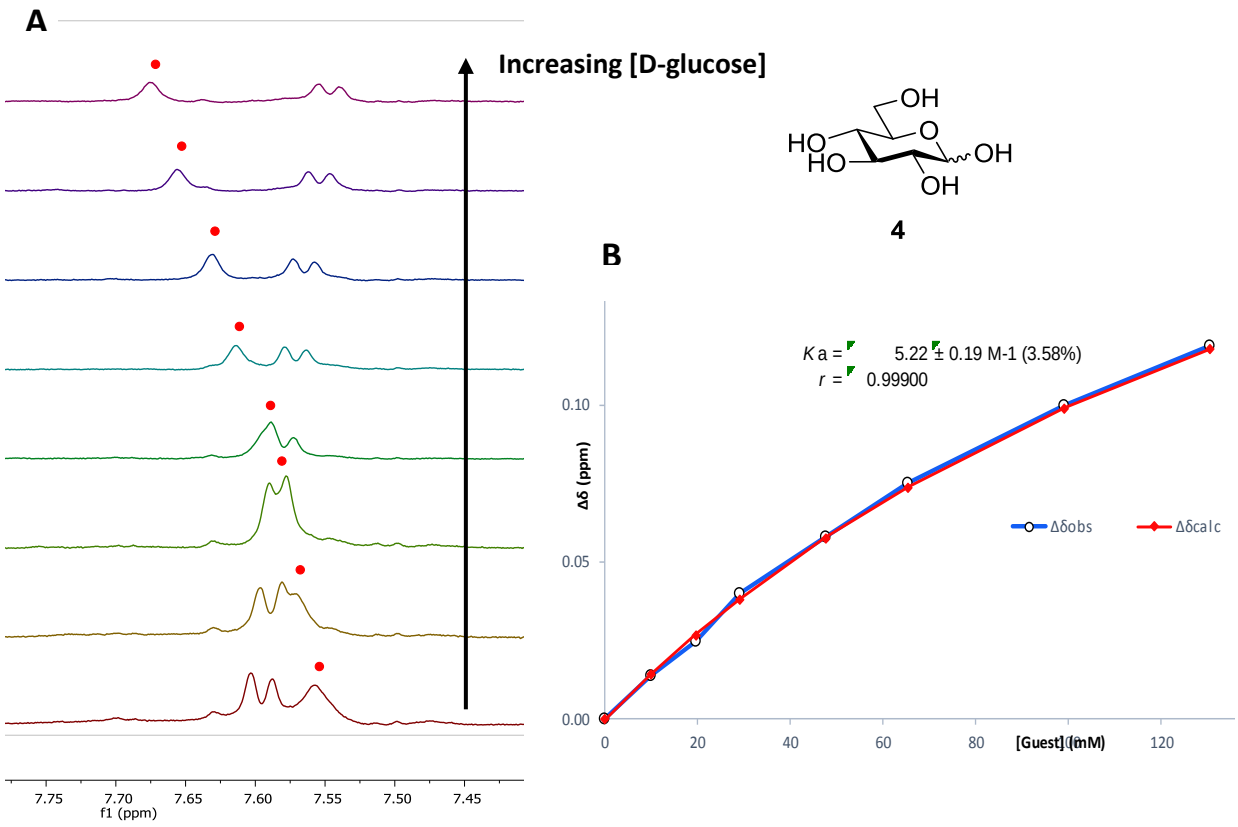


Figure 41 Partial ^{1}H NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) (A) and binding analysis curve (B) for receptor **1** (0.5 mM) titrated with a combined solution of D-glucose **4** (1 M) and receptor **1** (0.5 mM). Spectra imply binding with fast exchange on NMR timescale. Changes in chemical shift of peak δ 7.55 ppm (Δδ, denoted with •) were plotted against D-glucose concentration (mM). The calculated values for Δδ are overlaid with the observed values, giving K_a = 5.2 ± 0.2 M^{-1} (3.6%).

Upon addition of D-glucose **4** to receptor **1**, a change in chemical shift was observed for the two most resolved signals (relating to the spacer aromatic protons) in the ^{1}H NMR spectra. This would imply that binding was occurring with 'fast exchange' kinetics relative to the NMR timescale, which results in a time averaged spectrum of both the free host and host guest complex. This change in chemical shift (Δδ) was plotting against D-glucose concentration, and the curve fitted to a 1:1 binding model to yield an association constant (K_a), which afforded a K_a ~5 M^{-1} for receptor **1** and D-glucose **4**. This is approximately 12 times weaker than that observed for previous anthracene receptor **25**, and was rationalised to be due to the partially collapsed cavity of receptor **1**. Required disruption of the intramolecular interactions for binding of substrates to receptor **1** would lower the Gibb's free energy of binding (ΔG) and thus association constant (K_a), especially if said intramolecular interactions are particularly favourable. These interactions may not even be disrupted, with the D-glucose only partially encapsulated by the anthracenes and the two remaining ureas providing hydrogen bonding interactions – albeit in a less than ideal fashion. The amide spacer units and smaller distance between

anthracene surfaces prevent intramolecular hydrogen bonding of the amides and thus collapse of the cavity for previous anthracene receptor **25**, and this was believed to be the reason for the higher affinity observed for D-glucose for this receptor. The smaller cavity size for receptor **25** did disfavour larger substrates such as disaccharide D-cellobiose **31** however, which featured a K_a of ~28 M^{-1}.[35] NMR titration of receptor **1** with D-cellobiose **31** however showed an increased association constant of K_a ~ 50 M^{-1}, which was believed to be due to the larger cavity size of **1**. It is possible that receptor **1** is able to take advantage of the larger substrate size of D-cellobiose **31** (compared to D-glucose **4**), with an increased number of intermolecular interactions to give the observed increased affinity when compared to previous receptor **25**.

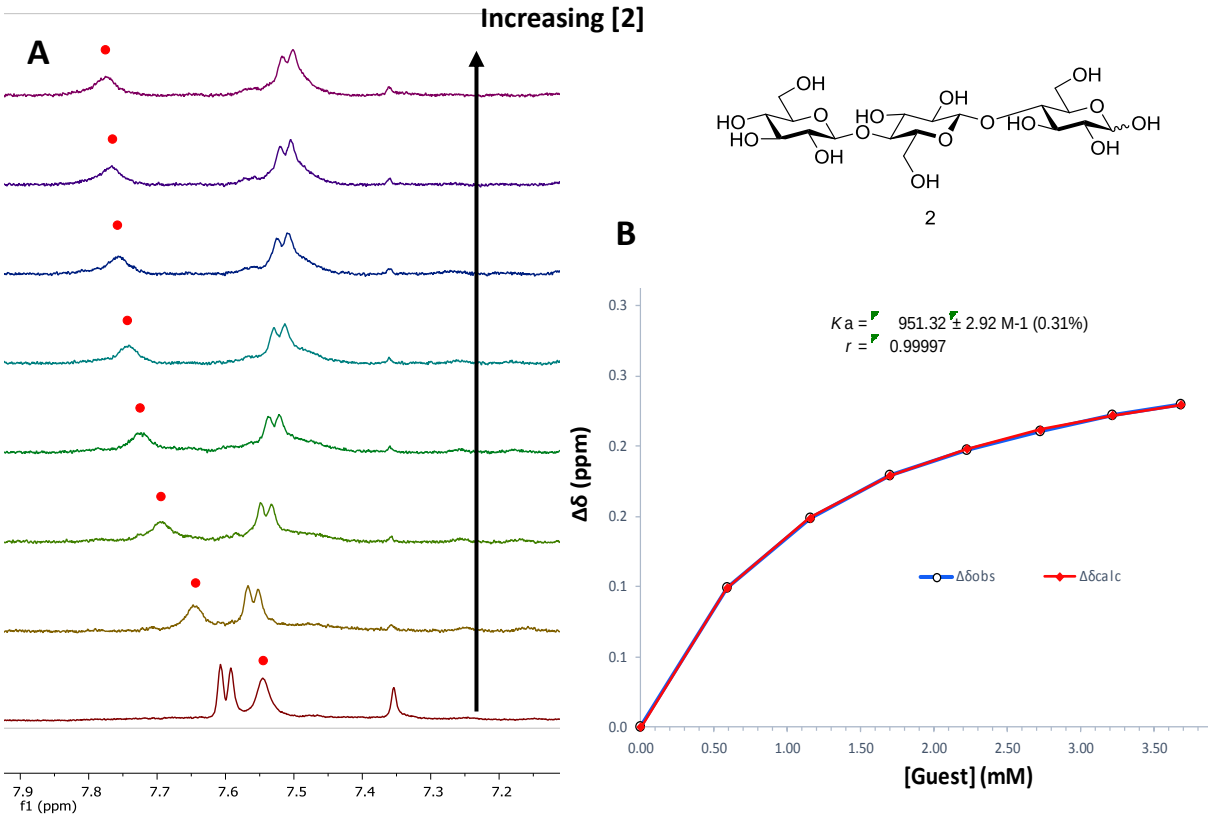


Figure 42 Partial ^{1}H NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) (A) and binding analysis curve (B) for receptor **1** (0.2 mM) titrated with a combined solution of D-cellotriose **2** (30 mM) and receptor **1** (0.2 mM). Spectra imply binding with fast exchange on NMR timescale. Changes in chemical shift of peak δ 7.55 ppm (Δδ, denoted with •) were plotted against concentration of **2** (mM). The calculated values for Δδ are overlaid with the observed values, giving K_a = 951 ± 3 M^{-1} (0.3%).

This trend of larger substrates giving stronger affinities was observed for the longer cellodextrins, where NMR titration of D-cellotriose **2** with receptor **1** gave a much-increased binding affinity of K_a ~ 950 M^{-1} – 19 times stronger than for D-cellobiose **31**. It was theorised that the dramatically increased affinity is due to the full length of the cavity becoming involved in binding, with the substrate threading through until optimal interactions are achieved between receptor and substrate. Some limited evidence for this was that during the NMR titration of D-cellotriose **2** (and other substrates) with receptor **1**, the broad signal for the terminal anthracene protons showed a Δδ of ~0.5 ppm – moving from δ ~6.7 to 7.2 ppm. This was potentially an indication that the receptor cavity opens upon binding, disrupting the CH-π interactions between anthracenes, allowing larger substrates (such as D-cellotriose **2**) to thread through, and thus resulting in the more downfield shift for the terminal anthracene protons (Figure 43).

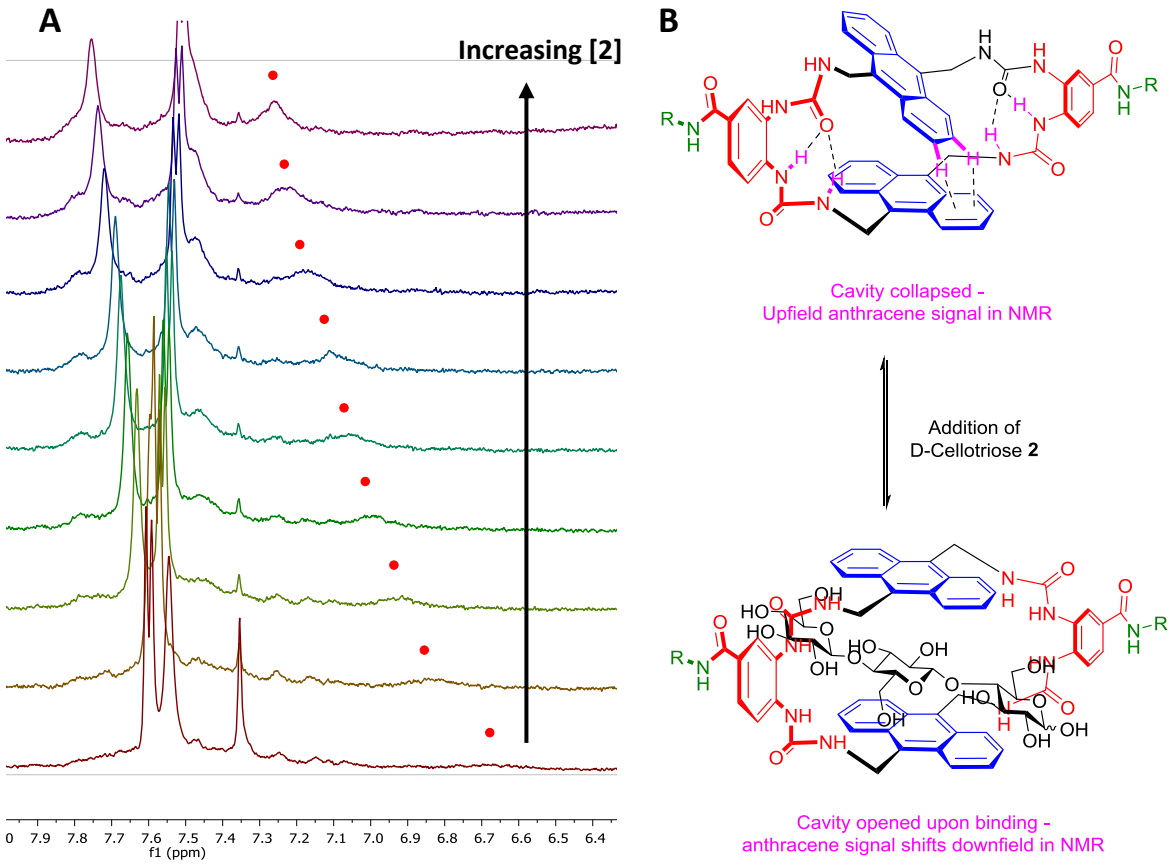


Figure 43 Partial ^{1}H spectra (600 MHz, D_2O, pH 7.4, 298 K) during NMR titration of receptor **1** and D-cellotriose **2**, showing the large change in chemical shift (Δδ ~ 0.5 ppm) for the terminal anthracene protons (A). This was theorised to occur due to opening of the anthracene cavity upon binding of the substrate, disrupting the CH-π interactions between anthracene surfaces and resulting in a more downfield chemical shift (B).

Attempts to ascertain a binding affinity for the larger cellodextrins: D-cellotetraose **33**, D-cellopentaose **130** and D-cellohexaose**131**, were unsuccessful by NMR titration. Severe broadening of

receptor signals were observed by NMR upon addition of each substrate during the titration – an indication that the binding kinetics were operating within an 'intermediate exchange' regime, meaning there was no way to extract any tangible data to fit to a binding model and thus calculate the association constant. The slower kinetics observed for these substrates could originate from the increased size of the substrates increasing the time taken to associate or dissociate from the receptor due to steric effects. However, this can also result from increased binding affinities increasing the time taken for the substrate to decomplex from the receptor, which would agree with the trend of increasing affinities with increasing cellodextrin length. Receptor **1** did appear to saturate at roughly a similar rate with these substrates as for D-cellotriose **2** however, which would qualitatively imply similarly strong affinities.

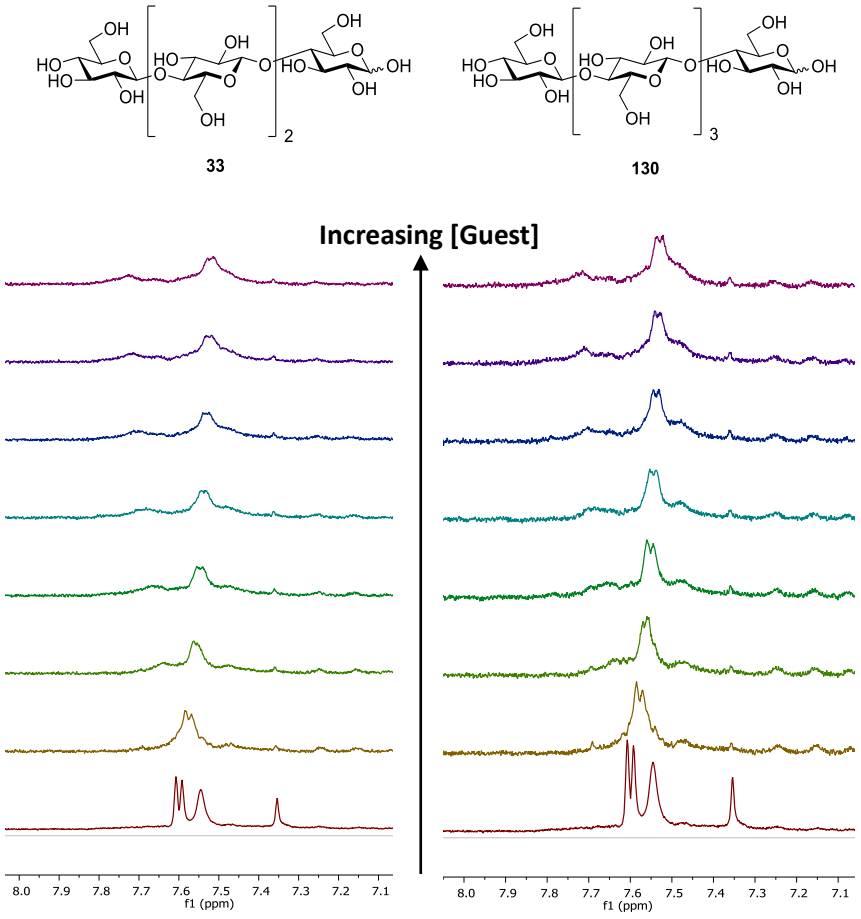


Figure 44 Partial ^{1}H NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) for NMR titration of D-cellotetraose **33** (left) and D-cellopentaose **130** (right) with receptor **1**. Spectra imply binding with intermediate exchange, whereby broadening of receptor signals were observed upon addition of substrate. Therefore, no K_a was determinable from these data. Some small peak movements were observed, but due to ambiguous chemical shift values, fitting of this data against guest concentration did not yield a binding curve presumably due to large errors in measured chemical shifts.

In order to determine the binding affinity for the longer cellodextrins, as well as corroborate the results already obtained, an alternate method was sought other than NMR titration. Therefore, binding affinities were investigated using isothermal titration calorimetry (ITC) experiments.[70] This technique titrates a guest solution into a solution of host and measures the heat changes upon each addition. This is done by the instrument maintaining the temperature of the sample cell (where the host solution resides) relative to a reference cell of pure water. The power applied by the instrument, to offset these changes in heat between the two cells, is then used to determine the exotherm upon addition of guest to host. Integration of the exotherm peak upon each addition gives the change in enthalpy (ΔH) for that point in the titration. Collation of this data and plotting the change in enthalpy versus concentration gives a curve – this can be fitted to a binding model (as for NMR titrations) to yield a binding constant (K_a). It is instructive to consider the enthalpy of dilution upon addition of guest to host solutions (i.e. enthalpy of dilution of sugars in water). Therefore, each ITC experiment consists of a 'blank' run where substrate is added to pure water (no receptor) and the enthalpy of dilution is obtained. Subtraction of this 'blank' from the titration of substrate into receptor effectively removes the heat of dilution from the titration – any enthalpy changes that remain must be a result of interaction of substrate with receptor (i.e. binding events).

Titration of D-cellobiose **31** (200 mM) into receptor **1** (0.2 mM) using ITC gave an association constant of K_a ~40 M^{-1}, comparable to the NMR titration results. ITC titration of D-cellotriose **2** (15 mM) into receptor **1** (0.2 mM) gave much a larger exotherm upon addition of substrate, and fitting this data to a 1:1 binding model yielded the expected higher affinity of K_a ~ 950 M^{-1} – essentially the same as for NMR titration. A distinct advantage of using ITC to determine binding affinities (compared to NMR for example) is that the thermodynamic properties of binding can be extracted from the ITC data. The enthalpy of binding (ΔH) is equivalent to the heat change per mol of host-guest complex formed. The binding affinity (K_a) can be used to determine the Gibbs free energy of binding (ΔG) using the equation: $\Delta G = RT\ln K_a$. This then allows determination of the entropy of binding (ΔS) using the equation: $\Delta G = \Delta H - T\Delta S$. Applying this to the ITC results obtained for D-cellotriose **2** and **1** shows that the main contribution of binding appears to be enthalpically based, with a much smaller contribution from the entropy of binding. This is most likely due to the strong intermolecular interactions formed between **2** and the receptor **1** being the main driving force for complexation – such as multiple CH-π interactions between the oligosaccharide CHs and the anthracene.. Large increases in entropy upon binding are usually indicative of many water molecules being displaced from the binding site or an increase in receptor/substrate flexibility upon binding.[71,72] Neither would appear likely in the case of **1**, given its rigid aromatic surfaces and the large size of the substrate **2**, and this is reflected in the thermodynamic results.

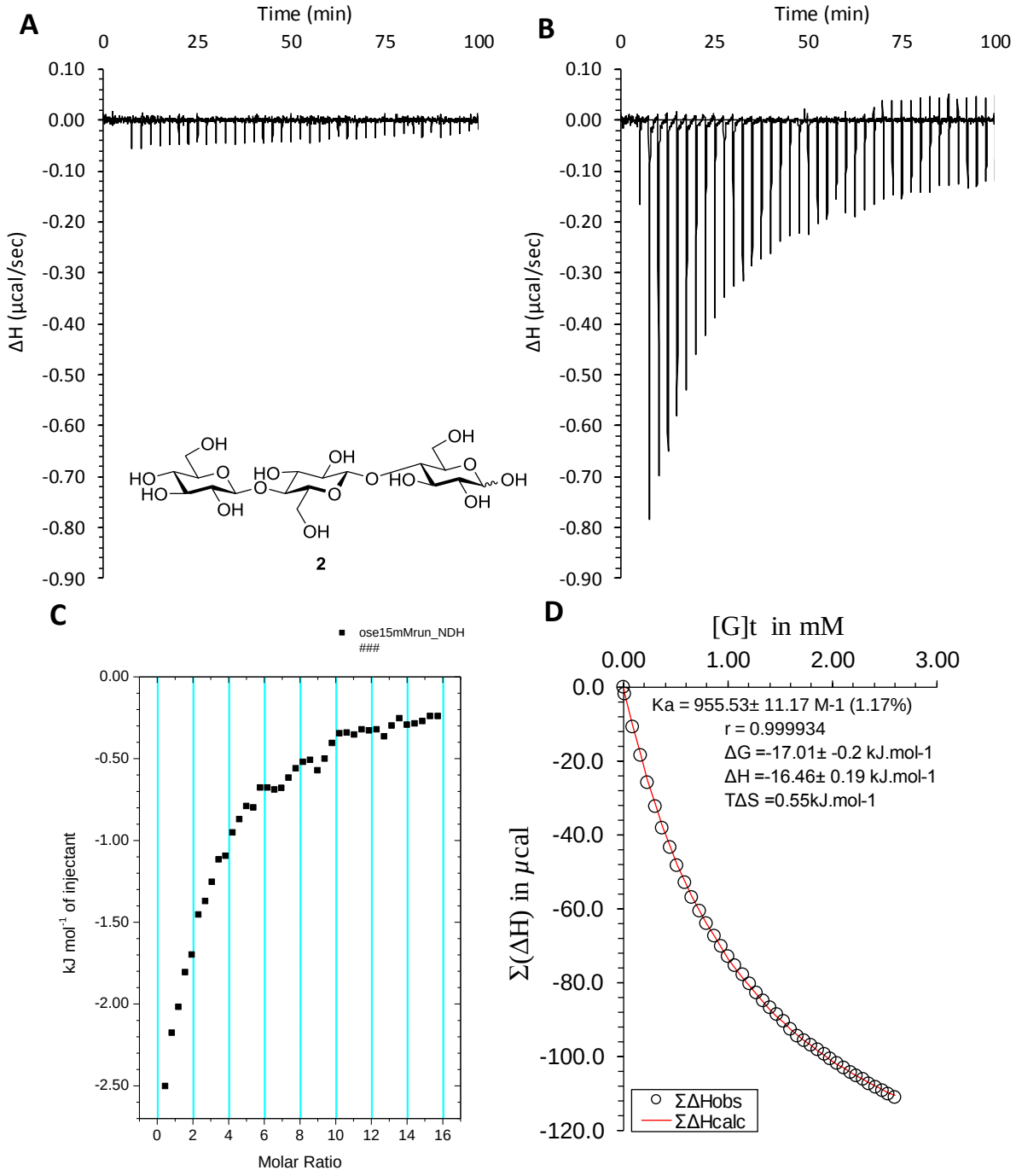


Figure 45 ITC binding results for receptor **1** (0.2 mM) titrated with D-cellotriose **2** (15 mM) in water (pH 7.4) at 298 K, in which: A) shows the blank run (addition of substrate into water); B) shows the titration (substrate into receptor); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 955 ± 11 M^{-1}).

Unfortunately, the affinities for the longer cellodextrins such as **33**, **130**, and **131** could not be determined using this technique (ITC) at the time of writing due to prolonged instrument failure. Therefore, further NMR studies were undertaken in the meantime to ascertain the selectivity of receptor **1** for linear oligosaccharides and whether threading of these substrates through the receptor

cavity was indeed occurring. NMR titrations were therefore performed with maltodextrins, which are α-1,4 linked oligosaccharides comprised of glucose monomers. These substrates are not linear, compared to the β-1,4 cellodextrins already tested, and are therefore unlikely to thread through the receptor cavity upon binding. If this was the case, then reduced affinities for these maltodextrins to receptor **85** (compared to the cellodextrins) are therefore expected. Indeed, such a result was obtained, where NMR titration of D- maltose **132** with receptor **1** yielded an association constant of K_a ~15 M^{-1} – approximately 3 times weaker than for D-cellobiose **31**. An even more supportive result was obtained from the NMR titration of D-maltotriose **133** with receptor **1** however, which afforded a binding affinity of K_a ~ 20 M^{-1} – only marginally higher than for D-maltose **132**. The affinity for non-linear trisaccharide **133** is approximately 48 times weaker than for linear D-cellotriose **2**, which would imply very different modes of binding to receptor **1**. This would strengthen the hypothesis that the linear cellodextrins are indeed threading through the receptor cavity, unlike the non-linear maltodextrins. Given the very similar association constants for D-maltose **132** and D-maltotriose **133** it could be reasonably speculated that only a small portion of the substrate (such as part of the terminal sugar) takes part in binding to receptor **1**, and that the steric effects of the α-linked glycosidic bonds make greater encapsulation by the receptor impossible.

Current work is underway to confirm the threading of the cellodextrin substrates through more in-depth NMR methods such as NOESY studies of the receptor-guest complexes. Given the intermediate exchange rate of binding for longer substrates (such as D-cellopentaose **130**), acquisition of these spectra at reduced or elevated temperatures is predicted to give more resolved signals pertaining to the receptor-guest complex by increasing or reducing the rate of exchange, and thus provide more accurate data. Critical signals that would indicate threading would be through space interactions (NOEs) between the receptor and central saccharide units or central glycosidic bonds.

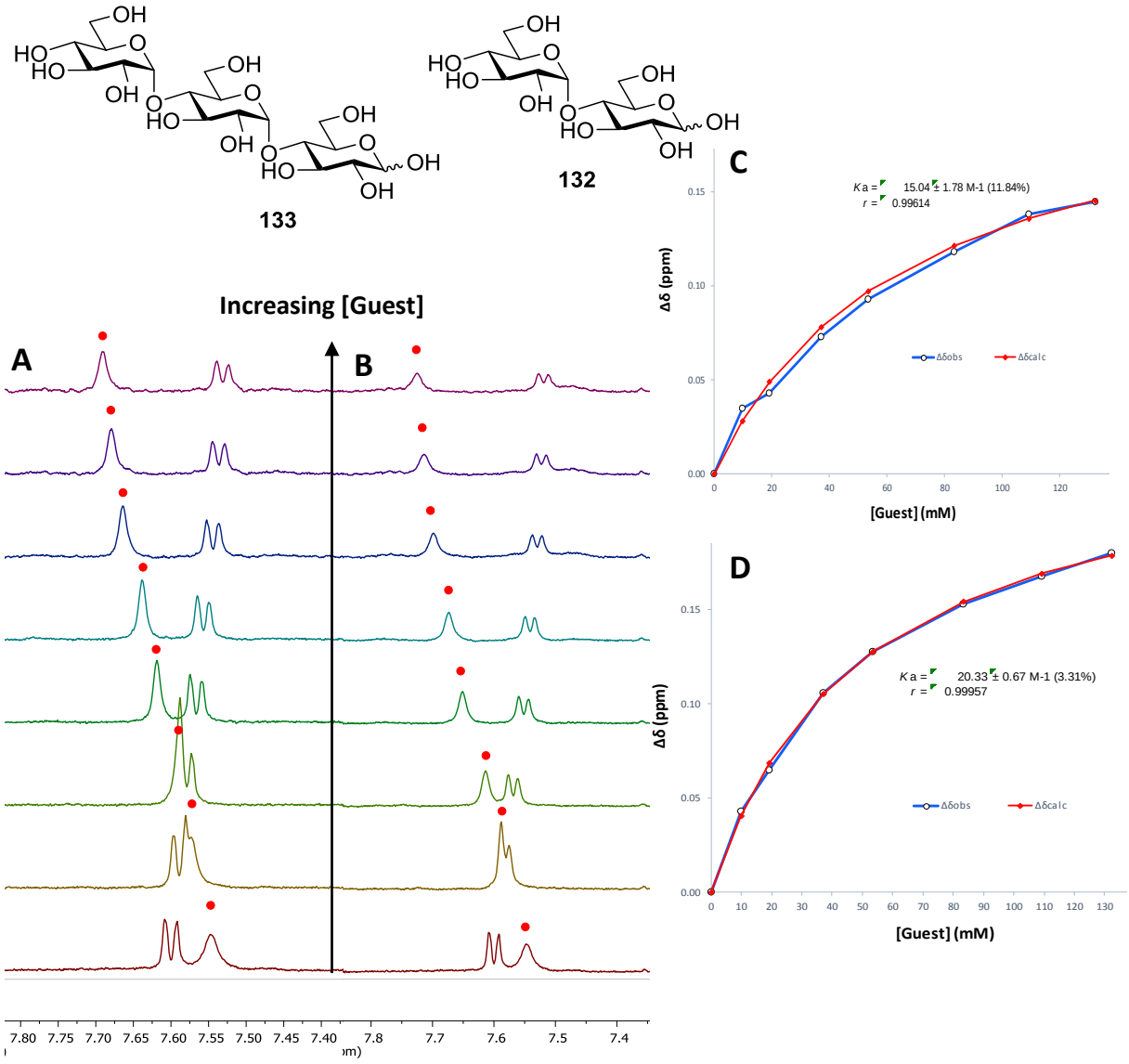


Figure 46 Partial ^{1}H NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) for titration of D-maltose **132** (A) and D-maltotriose **133** (B) with receptor **1** (0.2 mM). Spectra imply binding with fast exchange on NMR timescale. Changes in chemical shift of peak δ 7.55 ppm (Δδ, denoted with •) were plotted against concentration of **132** or **133** (mM). The calculated values for Δδ are overlaid with the observed values, giving K_a = 15 ± 2 M^{-1} (11.8%) for **132** (C) and K_a = 20 ± 0.8 M^{-1} (3.3%) for **133** (D).

Given the inherent fluorescent nature of anthracene, studies were performed to see whether there was a measurable fluorescence response upon binding of a substrate to receptor **1**, and thus whether a binding affinity could be obtained through fluorescence titrations. An excitation wavelength of 395 nm was used, which produced an emission maximum at 423 nm. This was plotted against guest concentration to yield the binding curves. Fitting this data to a 1:1 binding model yielded the association constants in a similar manner as for previously described methods (NMR and ITC). Fluorescence titrations of receptor **1** with D-cellobiose **31** showed an increase in fluorescence intensity

upon binding, with an overall 20% increase in emission intensity upon reaching saturation with guest. Plotting the change in emission intensity against concentration of **31** yielded a binding curve. Fitting this data to a 1:1 binding model yielded a K_a of ~50 M^{-1} – essentially the same as for NMR and ITC titrations. Fluorescence binding studies with D-glucose **4** gave an affinity of K_a ~ 5 M^{-1}, whereas methyl β-D-glucoside **14** gave a similar association constant of K_a ~6 M^{-1} with this method. The fact that **14** also features a very low affinity for receptor **1**, despite its increased hydrophobic nature due to methylation at the anomeric hydroxyl, would suggest that they bind in a similar manner. It would also suggest that only larger linear carbohydrates bind to receptor **1** with any appreciable affinity, as they are able to maximise interactions with the receptor cavity due to their increased length – presumably due to increased number of CH-π interactions between the axial CHs on the sugar and anthracene surfaces.

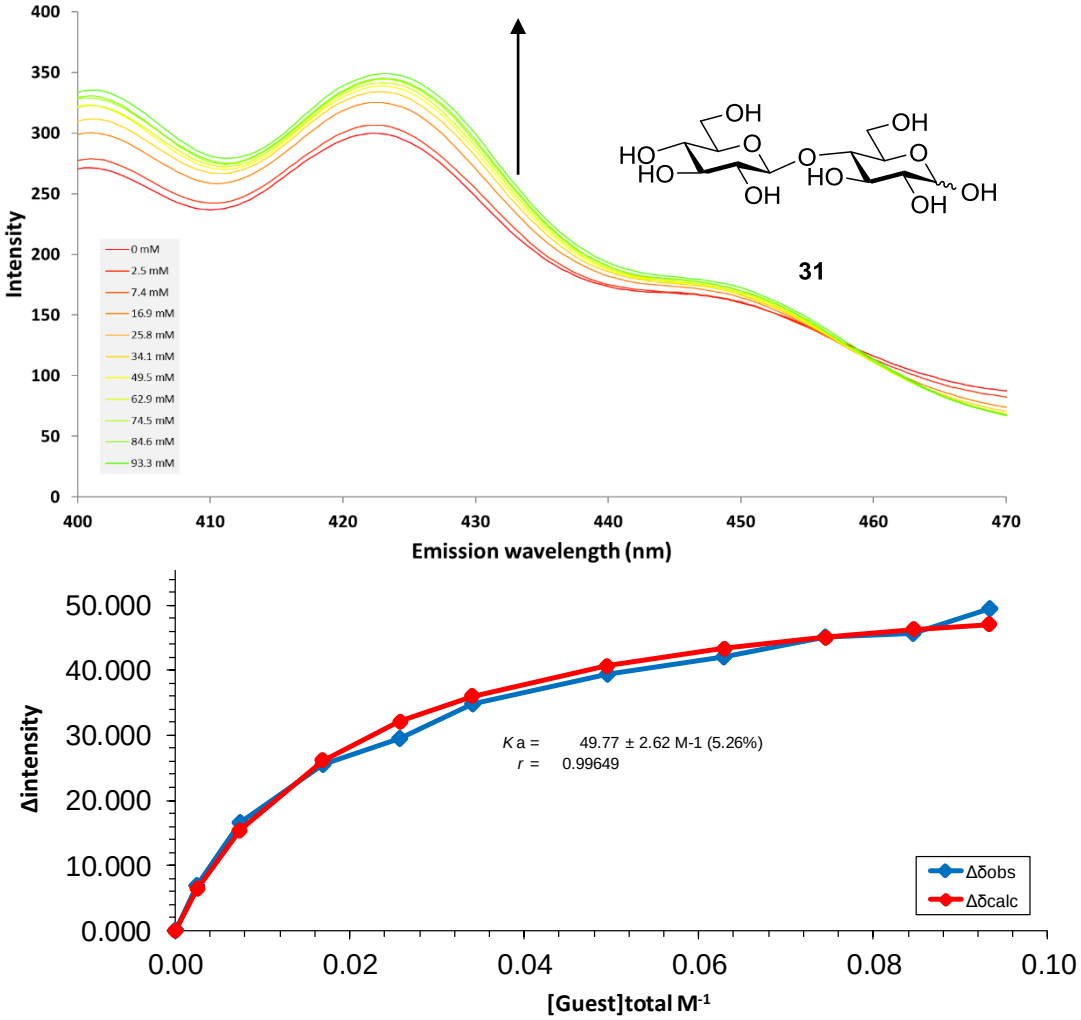
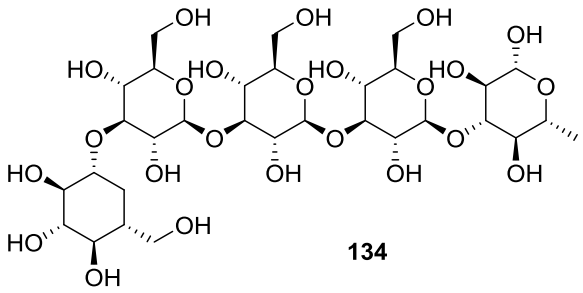


Figure 47 Fluorescence titration results for D-cellobiose **31** (150 mM) and receptor **1** (10 µM) in water (pH 7.4) at 298 K. Excitation wavelength of 395 nm used, while monitoring emission intensity at 423 nm. Changes in emission intensity at 423 nm were plotted against concentration of D-cellobiose **31**. The calculated values for change in intensity are overlaid with the observed values giving K_a = 50 ± 2.6 M^{-1} (5.2%).

An increase in fluorescence emission upon binding was also observed for previous anthracene receptor **25**, whereby close contact of the anthracene surfaces in the free receptor resulted in self-quenching of the fluorescence emission.[35] Binding of a guest to this receptor separated the anthracene surfaces and removed this quenching pathway, thereby increasing overall fluorescence emission intensity upon binding. Given that a similar result was observed for tetraurea anthracene **1**, and that previous NMR data suggests close contact between the anthracene surfaces in the free receptor, it would seem reasonable to assume a similar mechanism is the source of increased emission intensity for **1** upon binding as well.

Quantification of binding affinities for the longer cellodextrins (i.e. D-cellopentaose **130**) was not possible by fluorescence titration due the relatively large volume of the sample cell (3 mL), compared to NMR titration (0.5 mL). This results in needing a much larger quantity of guest needing to be added to drive the equilibrium of binding towards complexation, as guest titrant solutions are significantly diluted upon addition to the sample cell. This was not possible given the small quantities of oligosaccharides possessed. Future work is currently pursuing potentially the use smaller sample cells for fluorescence, as well as other techniques such as induced circular dichroism (ICD) to determine binding affinity for these longer cellodextrins, as this method was successfully deployed in determining binding constants for another receptor system (see chapter 5).[73] ICD also uses similar sample cell volumes to NMR titrations and would therefore be more applicable to investigating the binding properties of these longer cellodextrin substrates.

This measurable fluorescence output upon binding is desirable, as it creates the potential for receptor **1** to function as sensor. Given its affinity and apparent selectivity for linear cellodextrins, it could be applied towards the detection of β-glucans (such as **134**, Figure 48) which are a natural component of the cells walls of bacteria and fungi.[74,75] Receptor **1** could then be applied to a detection assay for invasive bacterial and fungal infections in biological samples, as detection of these characteristic oligosaccharides by **1** would be indicative of their presence, leading to a faster diagnosis.[76,77] Future work involving binding studies of **1** with β-glucan substrates will be necessary to further this potential application however.



134

Figure 48 Example structure of a typical β-glucan **134**, which features a linear oligomeric structure of D-glucose monomers linked with β-glycosidic bonds.

The binding results obtained for tetraurea anthracene **1** and all substrates tested using various techniques (NMR, ITC and fluorescence) are compiled in Table 4 below. When listed in this format, receptor **1** clearly shows a preference for binding linear glucose derived saccharides (i.e. cellodextrins), as very poor binding was observed for non-linear substrates of similar monomer lengths (i.e. maltodextrins). The weak affinities to smaller substrates suggest that they do not take full advantage of the receptor cavity, unlike longer substrates such as D-cellotriose **2** which showed almost a 50 times increased affinity over D-cellobiose **31**. Future work is currently underway to ascertain the binding affinities for the longer cellodextrins, as well as corroborate the other binding results with secondary methods. The binding interactions of receptor **1** with cellodextrins is also being investigated with exploratory NMR techniques such as NOESY, to ascertain whether the oligosaccharides are threading through the receptor cavity as predicted.

Table 4 Summary of binding results for receptor **1** with various substrates and their associated errors for each technique used (NMR, ITC or fluorescence). All binding studies conducted in H_2O (or D_2O for NMR), at pH 7.4 and at 298 K unless otherwise stated. [a] n.d. = not determinable due to intermediate exchange binding kinetics.

Substrate	**Binding constant K_a (M^{-1}) determined by:**		
	NMR	**ITC**	**Fluorescence**
D-Glucose	5 ± 1.0%	-	5 ± 6.2%
D-Cellobiose	46 ± 0.9%	38 ± 6.5%	49 ± 5.3%
D-Cellotriose	950 ± 0.3%	960 ± 1.2%	-
D-Cellotetraose	n.d.[a]	-	-
D-Cellopentaose	n.d.[a]	-	-
D-Cellohexaose	n.d.[a]	-	-
Methyl β-D-glucoside	-	-	6 ± 3.5%
D-Maltose	15 ± 11.8%	-	-
D-Maltotriose	20 ± 3.3%	-	-

While the magnitude of affinities obtained for **1** and oligosaccharide guests far exceed those of previous anthracene receptor **25**, they do not fully compare to previously reported pyrene receptor **32** – which featured an affinity 5 times as strong for D-cellotriose **2** of K_a ~ 5000 M^{-1}.[39] This was speculated to be due to the more rigid macrocyclic structure preventing collapse of the cavity, unlike **32** which can form intramolecular hydrogen bonds between ureas and CH-π interactions between the anthracene surfaces. However, pyrene receptor **32** was much more difficult to synthesise and so anthracene receptor **1** represents a much more accessible system that still features substrate affinities and selectivities comparable to natural carbohydrate binding proteins (such as lectins). Another advantage of **1** over **32** was that a fluorescence response was observed upon binding of guests to **1**, postulated to be due to conformational changes upon binding. No such fluorescence activity was observed for **32**, and this was attributed to the previously mentioned rigid structure. Pyrene receptor **32** also featured diminishing binding constants with increased cellodextrin length, where D-cellotetraose **33** gave the highest observed affinity but D-cellopentaose **130** and D-cellohexaose **131** then gave a reported 30% reduction in affinity compared to **33**. This was speculated to be due to the tripodal arrangement of the spacer units, where longer substrates (such as **130** and **131**) would then feature steric clash with one of the large solubilising groups. This would not be predicted for anthracene system **1**, as the solubilising groups are perpendicular to the anthracene surfaces and thus cannot interfere with the substrate once bound to the receptor. It would therefore be predicted that binding affinity to **1** would increase with substrate length, although if the oligosaccharide became long enough two receptors binding to one substrate could be possible. This would emulate the ultimate function of the receptor threading onto cellulose and thus assisting in solubilising the polymeric material in a neutral aqueous environment.

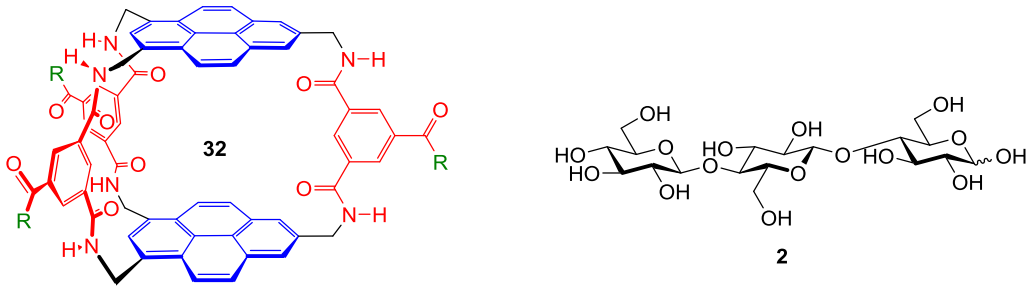
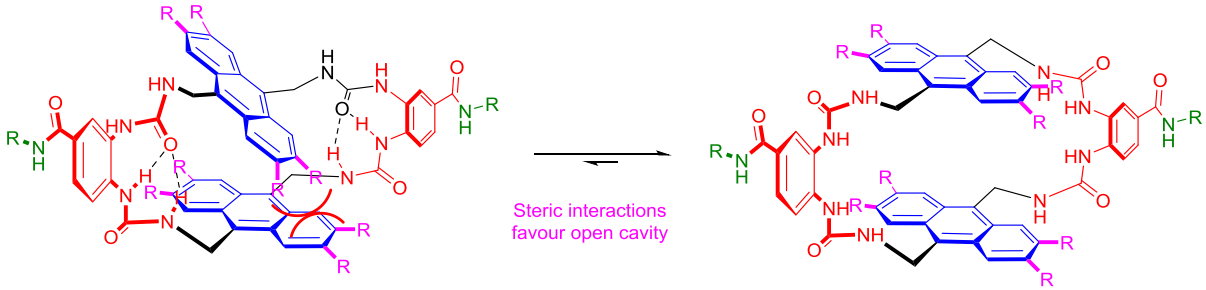


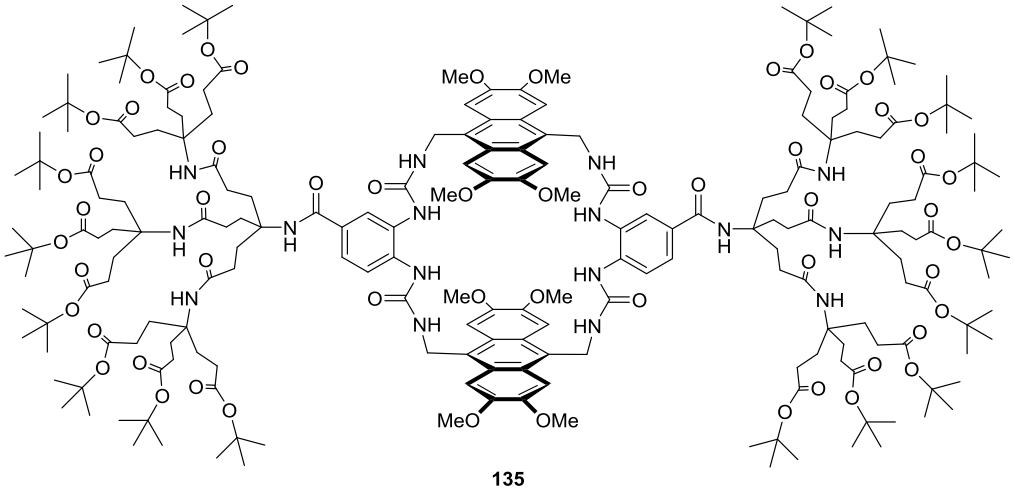
Figure 49 Previously synthesised pyrene based receptor **32** reported an association constant of K_a ~ 5000 M^{-1} for D-cellotriose **2**. While this affinity is markedly higher than for anthracene receptor **1**, **32** is much more difficult to synthesise than **1** and does not output a fluorescence response upon binding. R = G2MM solubilising group.

4.4 Receptor derivatives that extend the anthracene surface



Scheme 65 Schematic outlining the prediction that terminal substitution of the anthracenes would disfavour the collapsed cavity predicted for receptor **1**, due to unfavourable steric interactions. It was speculated that promotion of the open cavity conformation of the receptor would increase binding affinities. R = sterically bulky group.

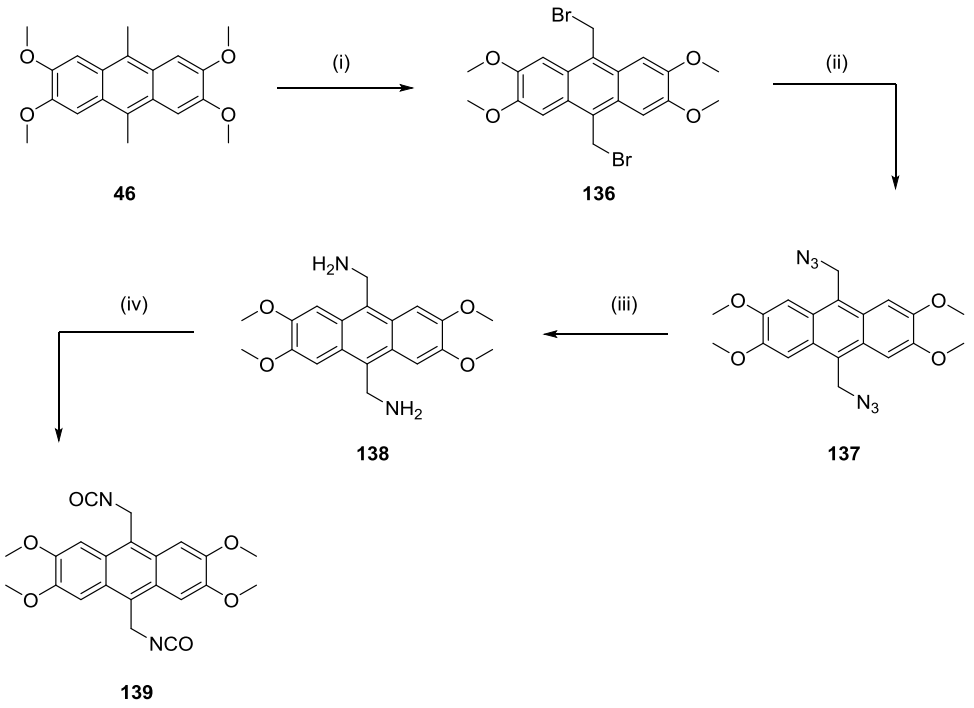
With the hypothesis that free receptor **1** adopts a conformation resulting in a partially collapsed cavity in water, and that this potentially diminishes binding affinities, receptor derivatives with extended anthracene surfaces were proposed. The increased length of the anthracene surfaces would be designed to promote a more open receptor conformation, through unfavourable steric interactions between anthracene surfaces forcing the anthracene surfaces to lie parallel to each other (Scheme 65). Therefore, receptor design **135** was proposed, which featured four methoxy groups at the terminal positions of the anthracene. These methoxy groups would increase the length of the anthracene surfaces but also increase the electron density of the anthracene ring system, which would strengthen the CH-π interactions between the anthracene and a carbohydrate substrate – resulting in increased binding affinities.



135

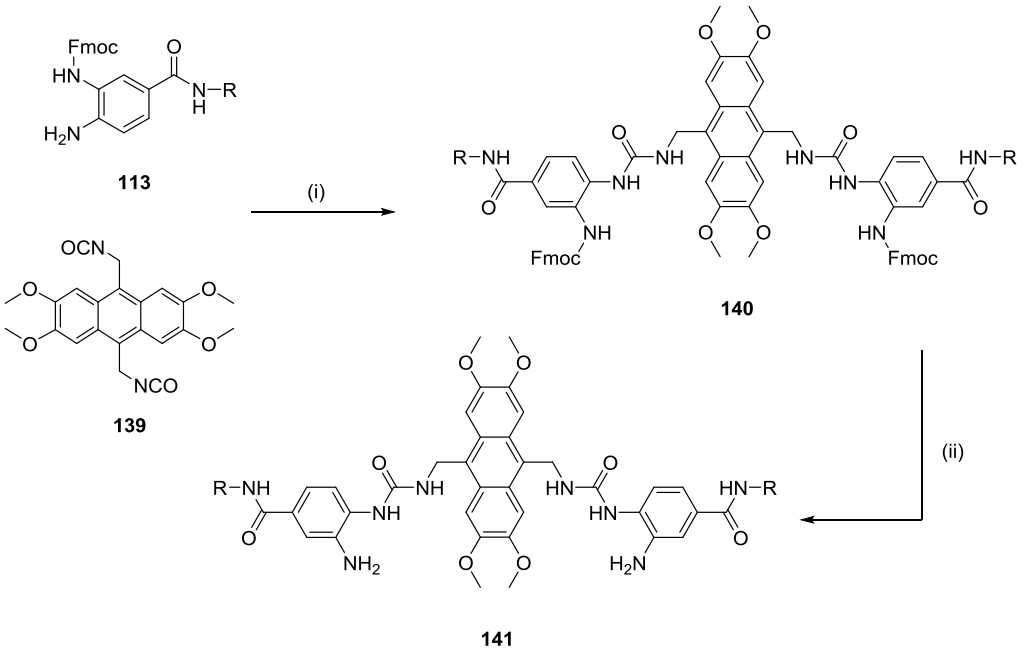
Figure 50 Proposed receptor design for protected anthracene macrocycle **135**, which substitutes the terminal positions of the anthracene with methoxy groups. These increase the length of the anthracene surface which disfavour collapse of the cavity (Scheme 65) through steric interactions. They also increase the electron density of the anthracene surfaces which are predicted to strengthen any CH-π interactions between the anthracenes and carbohydrate substrates.

Synthesis towards macrocycle **135** commenced through construction of the anthracene scaffold, which began with previously synthesised tetra-methoxy anthracene **46** (Scheme 66)**.** A radical bromination with NBS and radical initiator ABCN afforded dibromide **136** in good yield, followed by conversion to the diazido anthracene **137** in good yield *via* substitution with sodium azide. A Staudinger reduction using triphenylphosphine afforded diamine **138** in good yield, which was then converted to the diisocyanate **139** in moderate yield using triphosgene. Presence of the isocyanate functional group was confirmed *via* IR spectroscopy where the characteristic NCO stretch was observed at 2255 cm^{-1}. The isocyanate carbon was also observed with ^{13}C NMR spectroscopy.



Scheme 66 (i) NBS, ABCN, CH_2Cl_2, reflux, 4h, 71%; (ii) NaN_3, MeCN, reflux, 16h, 70%; (iii) PPh_3, THF, H_2O, 60 °C, 16h, 76%; (iv) triphosgene, toluene, reflux, 2h, 54%.

Formation of Fmoc protected half receptor **140** proceeded as before, whereby reaction of linker **113** with diisocyanate **139** in the presence of pyridine afforded **140** in good yield (Scheme 67). As for Fmoc protected half receptor **118**, **140** showed extreme broadening during characterisation by NMR but was confirmed by high resolution ESI mass spectrometry. Deprotection of the Fmoc groups with DBU proved trivial, affording amino half receptor **141** in near quantitative yield.

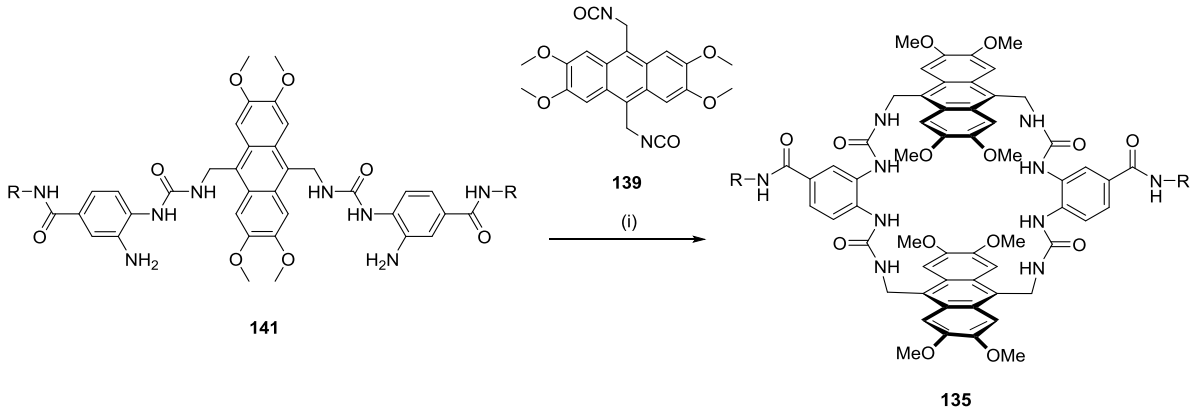


Scheme 67 (i) pyridine, CH_2Cl_2, reflux, 16h, 65%; (ii) DBU, CH_2Cl_2, 0 °C, 1h, 92%. R = G2MM solubilising group.

The methodology developed for forming protected anthracene macrocycles **120** and **126** proved completely transferable to this methoxy-anthracene system thus far. Indeed, formation of the octa-methoxy macrocycle **135** also proved comparable to previous systems when employing the optimised macrocyclisation conditions (Scheme 55, conditions ii). Reaction of amino half receptor **141** with diisocyanate **139**, in the presence of DMAP at a reaction concentration of 0.5mM in refluxed dichloromethane, afforded the target macrocycle **135** in a moderate yield of 32% (Scheme 68). Purification of the macrocycle was largely the same by reverse phase HPLC as for *tert*-butyl protected macrocycle **120**, eluting in acetone/water.

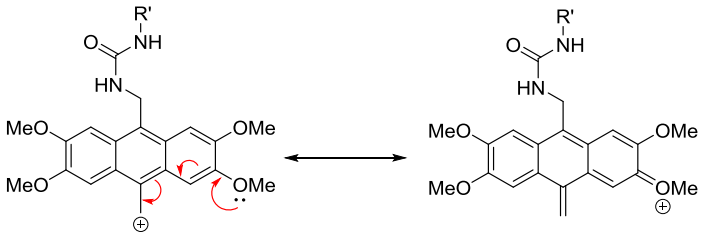
As for protected anthracene macrocycle **120**, the two different anthracene surfaces were visible (made inequivalent due to the asymmetric spacer unit) during characterisation with ^{1}H and ^{13}C NMR in CD_3OD. However, one notable difference was that signals pertaining to the anthracene surfaces were noticeably broader for **135** when compared to previous anthracene macrocycles **120** and **126**. This was theorised to be due to the increased sterics of the methoxy groups hindering rotation of the anthracene surfaces, resulting in slower conformational exchange and thus broader NMR signals. This was deemed promising, as the increased length of the anthracene units were already having some effect on potential conformation. However, whether this effect would persist in aqueous media was yet to be seen. Macrocycle **135** did show incredible susceptibility to oxidation with singlet oxygen however, presumably due to the more electron rich nature of the anthracene surface – akin to the oxygen sensitivity of tetra-alkoxy anthracene macrocycle **37** (chapter 2). Macrocycle **135** was observed to undergo rapid oxidation in organic solvents during NMR characterisation in a matter of days.

Therefore, acquisition of the water-soluble macrocycle was now critical in order to suppress this oxidation.

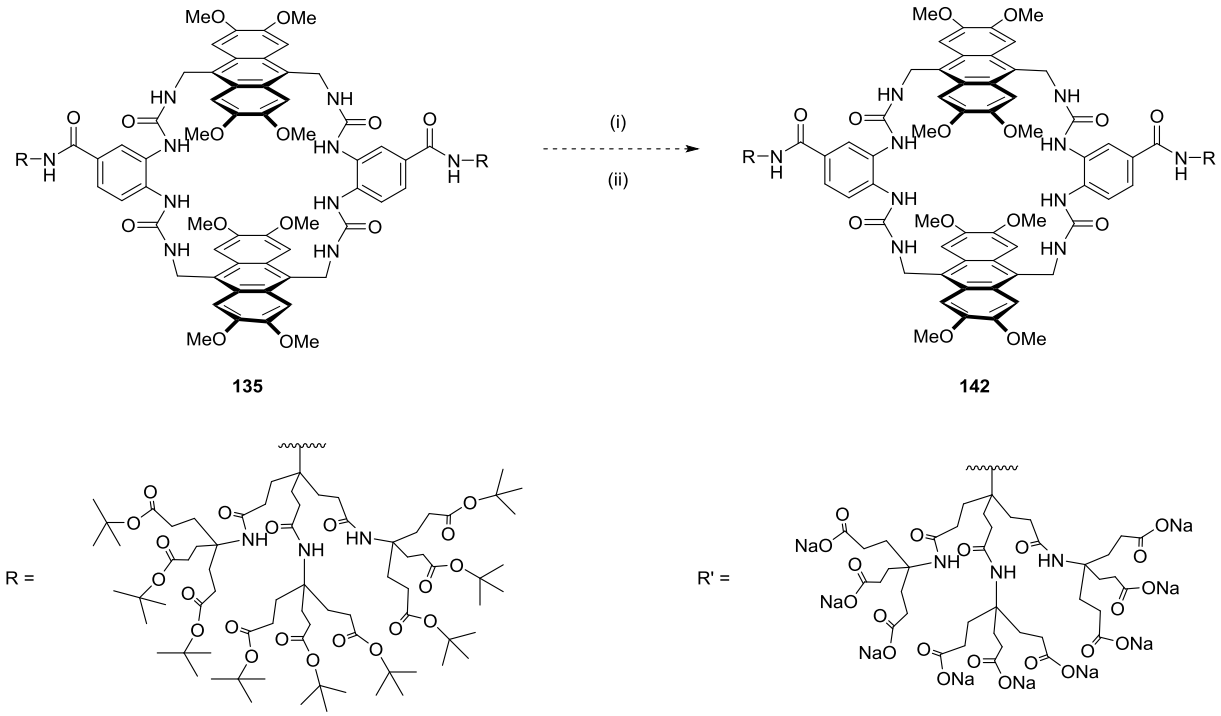


Scheme 68 (i) DMAP, CH_2Cl_2, reflux, 16h, 32%. R = G2MM solubilising group.

Unfortunately, exposure of protected macrocycle **135** to previously optimised deprotection conditions with 10% TFA (v/v) did not afford deprotected macrocycle **142 (**Scheme 70). Upon addition of TFA to tetraurea **135** a bright pink solution was observed. It was theorised that this was due to formation of a benzylic cation intermediate (described in chapter 2 and earlier in this chapter) through acid promoted heterolytic cleavage of the ureas. The extreme colour change was postulated to be due to conjugation of the mesomeric methoxy groups, stabilising the benzylic cation by delocalisation into the anthracene ring system (Scheme 69). The reaction was therefore stopped earlier than previous systems and attempted purification by reverse phase HPLC was undertaken. Most evident, however, was the presence of the spacer unit cleaved from the anthracene – as the major product visible by HPLC showed UV absorption at 210 and 254 nm but not 370 nm, which was typically used to monitor for the presence of anthracene. A very complex mixture of products was observed when monitoring at 370 nm however, presumed to be various decomposition products of anthracene.



Scheme 69 Proposed benzylic cation intermediate produced during the acidic deprotection of **135**.



Scheme 70 (i) TFA (10% v/v), CH_2Cl_2, rt, 2h; (ii) NaOH (aq) to pH 7.4, macrocycle **142** not observed. R = *tert*-butyl protected G2MM solubilising group. R' = nonacarboxylate G2MM solubilising group.

Given how notoriously sensitive the previous anthracene systems were to deprotection with acid, and that this octa-methoxy anthracene appeared even more unstable, it was decided to not pursue the acquisition of **142** any further at the time of writing. Future work could investigate the synthesis of anthracene surfaces that are much less electron rich such as tetra-carboxyl **143** or imide **144**, as these provide the desired increased surface length but also possess increased stability towards acid and oxidation. Alternatively, a long polyaromatic surface not derived from anthracene could be investigated, such as tetraoxa-pentacene **145**, which would be impervious to oxidation. The stability to acidic conditions of a tetraurea macrocycle incorporating **145** is not certain at this time however, but it would be predicted to be much more stable than **135** at the very least. The synthetic routes towards these three aromatic surfaces and their respective benzylic diamines is already known from previous work in the Davis group. Therefore synthesis of the corresponding macrocycles, using the optimised methodology discussed earlier in this chapter, would be predicted to be relatively straightforward. If acquisition of the water-soluble macrocycles can then be achieved, comparison of their activities to anthracene receptor **1** for cellodextrin substrates would be most interesting – especially to see whether extension of the aromatic surfaces promotes any positive effects towards binding affinities.

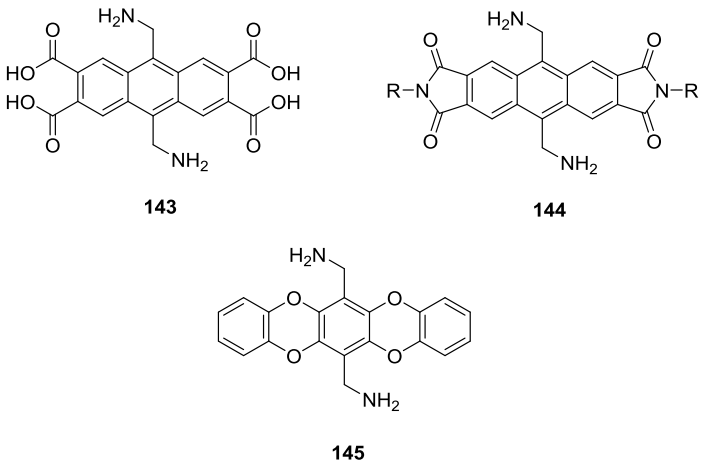


Figure 51 Potential aromatic surfaces that could incorporated into a tetraurea receptor. These all feature elongated aromatic surfaces compared to unsubstituted anthracene, and could provide increased binding affinities by promoting an open cavity conformation. The reduced electron density for each surface is also believed to be beneficial, by increasing stability to acid and oxidation. The synthetic route to each diamine is already known and thus their incorporation into the corresponding macrocycles is predicted to be feasible.

4.5 Conclusions

The urea spacer was successfully implemented into anthracene receptor **1**, which was synthesised from diisocyanate **60** and linker **113** in 4 steps, in 25% yield. Key optimisation was achieved in the deprotection of the *tert*-butyl protected solubilising groups with TFA, with new methodology enabling purification of the deprotected receptor **1** with reverse phase HPLC. Receptor **1** exhibited indications of intramolecular association between the ureas and anthracenes for the free receptor in water, with current work further characterising the conformation of the free receptor with NMR studies such as NOESY.

Receptor **1** showed good affinities for linear cellodextrins, such as D-cellotriose **2**, in water with affinities of K_a ~1000 M^{-1} – much increased over previous anthracene receptor **25** which was predicted to have too small a cavity to bind larger substrates such as cellodextrins.[35] The affinity shown by **1** is also very comparable to natural lectins.[19,78] The increased affinities here are attributed to the urea spacer units enlarging the cavity and providing an increased number of hydrogen bonding interactions. Selectivity over other oligosaccharides was good where tested, with linear cellodextrins binding ~50 times stronger than non-linear maltodextrins of the same monomer length. Work is currently underway to fully ascertain the binding affinities of longer cellodextrins (such as D-cellopentaose **130**), as well as confirming the selectivities over the weaker binding substrates. Full characterisation of the binding interactions between **1** and cellodextrin substrates with NMR is also underway and will confirm the hypothesis that the oligosaccharides are threading through the cavity. Once threading of receptor

1 onto oligosaccharides is confirmed, studies with larger fragments of D-cellulose **8** will be performed to ascertain whether threading is still taking place. Attempts can then be undertaken to see whether receptor **1** can assist in dissolving the polysaccharide in neutral water.

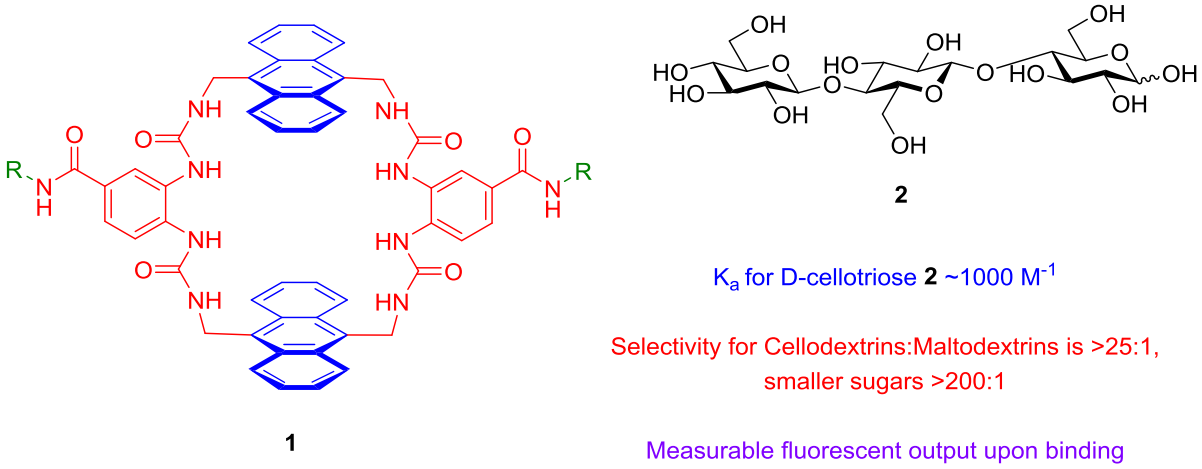


Figure 52 Anthracene tetraurea **1** was successfully synthesised and showed a distinct preference for the binding of linear cellodextrins, such as D-cellotriose **2**, which were bound selectively over monosaccharides and non-linear oligosaccharides (i.e. maltodextrins). Receptor **1** also produced a measurable fluorescent output upon binding, which could enable **1** to be employed as a sensor for the detection of β-glucans.

While receptor **1** did not match the binding performance of previous pyrene receptor **32**, it is far more synthetically accessible.[39] Tetraurea **1** also produces a measurable fluorescence response upon binding substrates. This, coupled with its affinity and selectivity for cellodextrins, make it a potential candidate as a sensor for the detection of β-glucans, which are characteristic biomarkers for invasive bacterial and fungal infections – these employ linear cellodextrins as components of the cell walls and are expected to be detectable by receptor **1.** Future study will be needed to confirm this through binding studies to β-glucans however.

With the modular synthetic approach towards receptor **1**, and seemingly robust synthetic methodology established, future synthetic work surrounding derivatives of **1** is predicted to be rich. The first point of derivatisation would be variation of the anthracene surfaces. Elongation of the surfaces could prevent cavity collapse and thus increase binding affinities. Altering of the electronic properties through electron donating or withdrawing groups could alter affinities, increase receptor stability and tune the fluorescence output.

Chapter 5 – Triethyl Benzene Hexaurea Receptor

5.1 D-glucose as a target substrate and towards treatments for Diabetes Mellitus

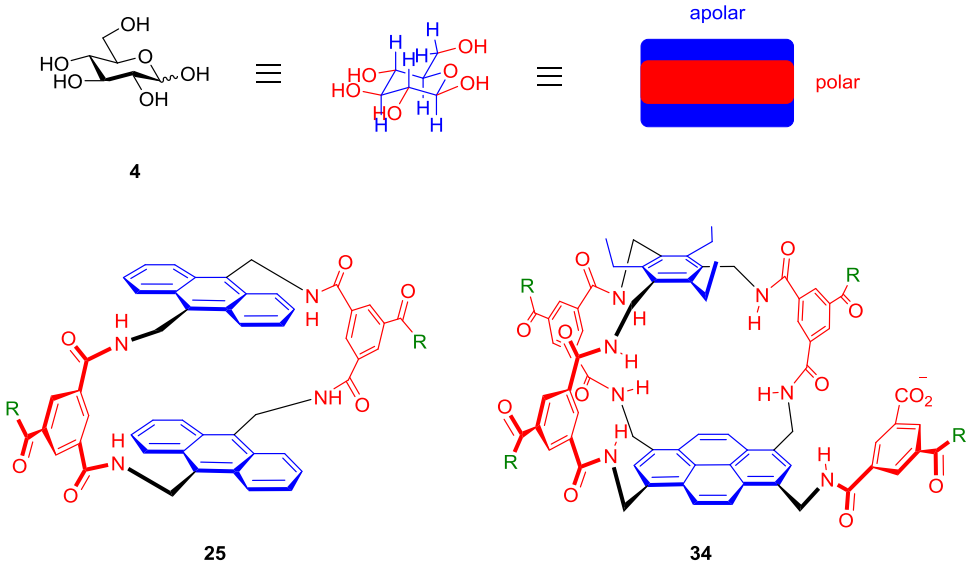


Figure 53 Common carbohydrate D-glucose **4** can be considered as featuring two distinct regions to exploit for binding interactions with a synthetic receptor: an apolar region on top and bottom comprised of CHs and a polar circumference of hydroxyls. Synthetic lectins **25** and **34** take advantage of these regions to afford affinities in water of K_a ~60 M^{-1} and 250 M^{-1} respectively.[35,40]

The carbohydrate D-glucose **4** is one of the most common organic molecules found in nature, and functions as a source of energy for most living organisms.[79] The most common isomer in aqueous solution is the pyranose (6-membered ring) form, which exists simultaneously as both α and β anomers (approximately 36:64, α:β ratio). Synthetic receptors looking to target D-glucose as a substrate will then focus on forming interactions with the pyranose form as it is the most abundant form in aqueous solution (>99%). Carbohydrate binding proteins (such as the lectin concanavalin A) traditionally target the α-anomer, as they can exploit the axial orientation of the anomeric hydroxyl groups with complementary binding motifs.[11] However, other proteins, such as hexokinases, can target the β-anomer exclusively.[80] It would be difficult to design a synthetic receptor that could bind both anomers with equal affinity, so traditional approaches to binding D-glucose target a single anomer. The β-anomer of D-glucose in particular features two distinct apolar and polar regions which can be exploited: axial CHs and an all equatorial arrangement of hydroxyls (Figure 53).[81] Previously reported synthetic lectins **25** and **34** displayed unprecedented (albeit modest compared to natural systems) affinities for D-glucose in water, through non-covalent apolar CH-π and polar hydrogen bonding interactions, with a preference for the β-anomer.[35,40]

D-glucose is of particular interest as a substrate for synthetic receptors as it plays a pivotal role in the disease Diabetes Mellitus (commonly referred to as just diabetes). Problems resulting in regulation of blood glucose levels give rise to diabetes, which causes very high blood glucose concentrations (hyperglycaemia). Long term symptoms of diabetes include kidney failure, blindness, brain damage and even death.[82] Diagnosis of diabetes is ever increasing in developed countries, with one in three people currently diagnosed with pre-diabetes (high risk of developing diabetes) in the United Kingdom. In all forms of diabetes, the body either does not produce the hormone insulin, or it no longer functions correctly, thus allowing blood glucose levels to remain unregulated. People suffering from diabetes must therefore manually monitor their blood glucose levels by drawing blood from their body (usually from a finger) and analysing the concentration of glucose in their blood. If the levels of glucose are too high, they must then manually administer insulin by injection to reduce their blood glucose levels. These readings are painful, not particularly accurate and must be performed several times a day – a significant practical and social hurdle to daily management of diabetes.

Future advanced treatments for diabetes are focused on developing continuous glucose monitors (smart sensors) – devices which would supply continual accurate blood glucose level measurements without the need to draw blood.[15,83] Also being developed in conjunction with these smart sensors are glucose responsive insulins (or smart insulins), which would independently recognise the presence of glucose in the blood and automatically release insulin only when it is needed, such as after a meal.[84] These treatments would effectively mimic the functions of the pancreas – leading to creation of an 'artificial pancreas' for diabetics, which would monitor and regulate their blood glucose levels autonomously.[85]

However, development of these idealistic treatments has been hindered by access to a suitable receptor that can selectively recognise D-glucose in a biological context (i.e. the blood). Most work has focussed on using natural proteins, such as the lectin concanavalin A and glucose oxidase, or boronic acid based receptors.[84,86,87] Both of these approaches suffer from numerous drawbacks including poor selectivity for D-glucose, pH sensitivity, lack of stability and toxicity.[88,89] Therefore, design and synthesis of a synthetic receptor that could selectively target D-glucose would be of significant interest, and could potentially revolutionise future diabetes treatments.

5.2 Receptor design and retrosynthetic analysis

Upon the synthesis of linker unit **113** (Chapter 3) and its successful incorporation into anthracene tetraurea receptor **1** (Chapter 4), efforts were directed towards implementation of **113** with other receptor designs. Previous synthetic lectin research in the Davis group has mainly focused on variation of the aromatic surfaces to generate new receptors. This required the synthesis of a variety of aromatic

surfaces with pendant amines (Figure 54).[33,35,38,39,60] Conversion of these amines to the isocyanate would allow reaction with linker **113** to yield the corresponding urea linked receptor derivatives. This would allow direct comparisons to be made with previously published receptors and better understand the effects these new urea linkages have on carbohydrate binding affinities.

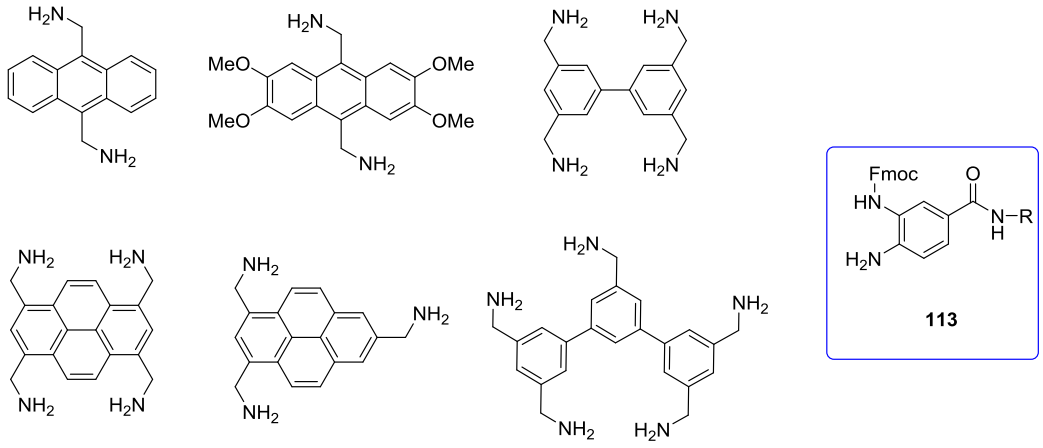


Figure 54 Aromatic amines previously synthesised in the Davis group that have been incorporated into various carbohydrate receptors. Conversion of these amines to isocyanates and reaction with linker **113** yields the equivalent urea receptors. R = G2MM solubilising group.[33,35,38,39,60]

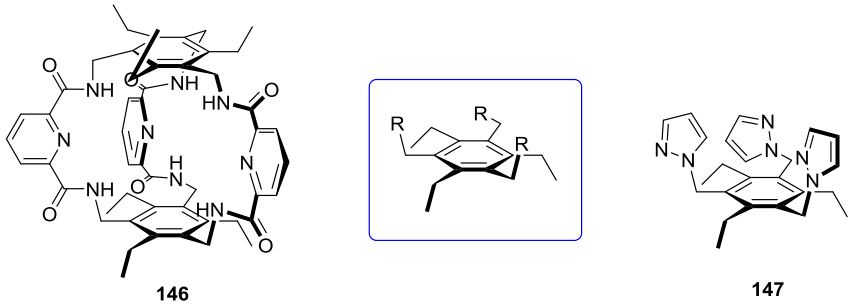


Figure 55 Incorporation of 1,3,5-substituted triethyl benzenes (blue box) into anion receptor **146** and cation binder **147**, as reported by Anslyn and co-workers.[90,91]

However, recent interest has turned towards utilising the popular supramolecular scaffold: 1,3,5-substituted triethyl benzene (TEB). This scaffold was initially pioneered by Anslyn *et al.*, where it was successfully incorporated into a variety of cation and anion receptors (Figure 55).[90,91] This scaffold has already seen some application towards carbohydrate recognition in organic solvents, with reports of macrocyclic receptor **16** by Roelens and colleagues.[23] Extensive investigation of carbohydrate recognition in organic media with derivatives of this scaffold has also been reported by Mazik *et al.*, with both acyclic **17** and monocyclic variants **18** showing appreciable activity towards their respective guests.[24,25] The TEB scaffold features steric preorganisation, in the case of acyclic architectures **147** and

17, where the ethyl and nitrogen functionalised groups (in these examples) tend to occupy the opposite face of the central aromatic ring.[92] Such preorganisation would be less integral in the case of closed cage structures like **16** and **18** but may assist in the favouring the conformation of previous intermediates required to form the macrocycle. Such steric effects may even lend to increasing the rigidity of the cage structure to maintain an open cavity.

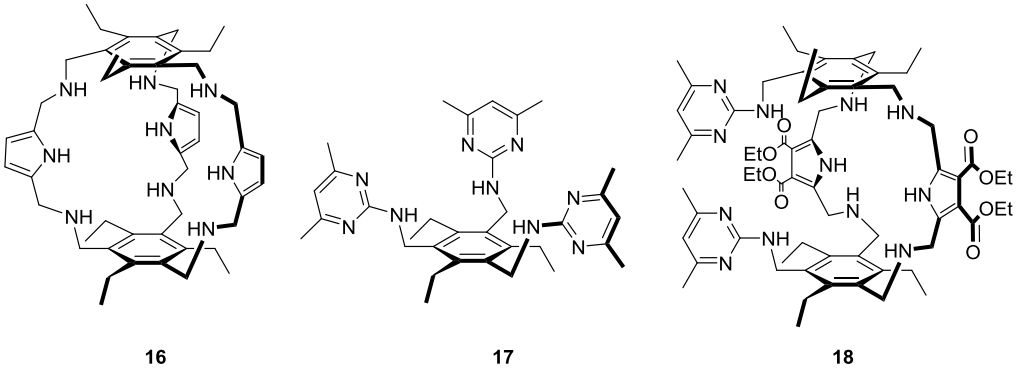


Figure 56 Examples of previous supramolecular architectures derived from triethyl benzene (TEB) scaffold that target carbohydrates in organic media. Pyrrole based macrocycle **16** reported by Roelens *et al*, and acyclic and monocyclic receptors **17** and **18** reported by Mazik and co-workers.[23,24,25]

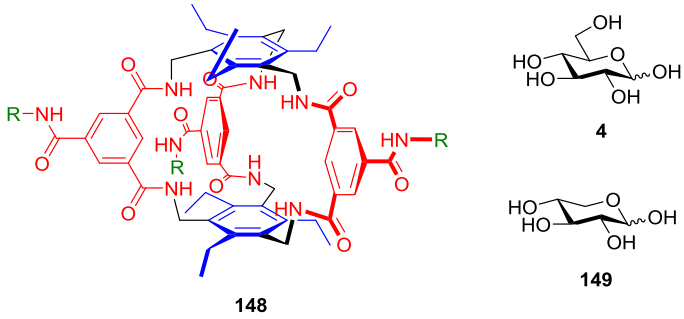


Figure 57 Unpublished results from the Davis group report synthesis of macrocyclic receptor **148**, which features the two TEB surfaces and three isophthalamide spacer units. Unfortunately, no binding to either D-glucose **4** or D-xylose **149** was observed by NMR binding studies, and this was believed to be due to the cavity being too small to accommodate the substrates. This hypothesis was corroborated by molecular modelling studies of receptor **148** and both guests inside the cavity. R= water solubilising group.[93]

The TEB scaffold has never been successfully incorporated into a receptor targeting carbohydrates in water however. Unpublished results from the Davis group outline the synthesis of hexa-amide **148**, which makes use of the previously discussed isophthalamide spacer units. Unfortunately, no evidence of binding to D-glucose **4** was observed by NMR spectroscopy binding studies. Subsequent molecular modelling then suggested that the cavity was too small, with the spacer units providing insufficient space between the aromatic surfaces for the substrate to fit. No binding was also witnessed with the

smaller sugar D-xylose **149**, which corroborates with the hypothesis that the cavity is indeed too small.[93] At the time of this work, no viable synthetic route to a larger spacer was known. However, with the successful synthesis of linker unit **113** (Chapter 3) and incorporation into anthracene tetraurea receptor **1** (Chapter 4), application of this linker unit **113** to the TEB scaffold could yield a receptor cavity of sufficient size to accommodate a desired carbohydrate guest. Thus, the design of receptor **3** was proposed, which would feature three bis-urea linkages (one more than anthracene receptor **1**, Chapter 4). The optimised design employs six ureas and thus up to twelve hydrogen bond donors directed towards the central cavity. Water solubility would be enabled by three pendant solubilising groups – each composed of G2MM dendritic nonacarboxylates, same as for anthracene receptor **1**. Given that anthracene receptor **1** was fully soluble with two solubilising groups and featured larger hydrophobic surfaces, it was predicted that the three G2MM dendrimers will provide ample water solubility for **3**, as well as preventing aggregation through electrostatic repulsion.

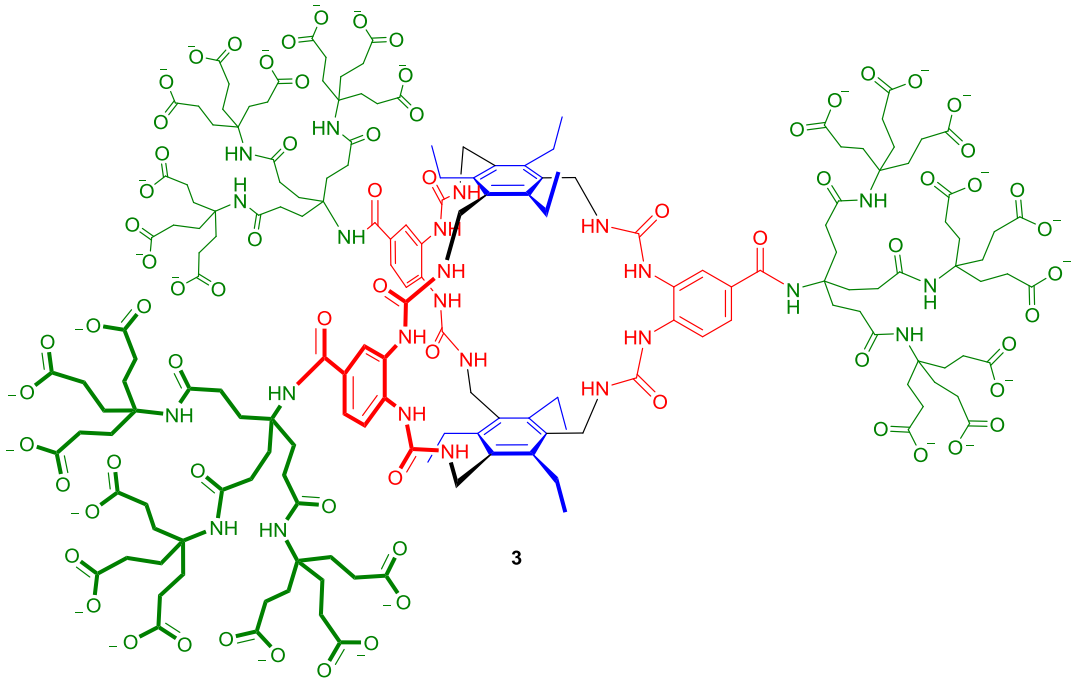


Figure 58 Proposed design for receptor **3**, which features six ureas to provide twelve hydrogen bond donors directed towards the central cavity. Initial attempts will make use of the G2MM solubilising group, as used for anthracene receptor **1** (Chapter 4).

The cavity of **3** would be smaller in size when compared to previous anthracene systems, but was predicted to be very complementary and thus selective towards all equatorial monosaccharides such as β-D-glucose **4**. Indeed, such predictions were corroborated by computational studies and molecular modelling of both the free receptor and receptor bound with various guests. MCMM conformational searches yielded a ground state conformation where **3** featured an open cavity that had not collapsed

or distorted. No intramolecular hydrogen bonding was observed between urea spacer units, in fact all ureas were predicted to be directed inwards towards the central cavity. Introduction of β-D-glucose into the cavity and modelling of host-guest interactions predicted that five out of six ureas participate in intermolecular hydrogen bonding. Only one urea of **3** and the hydroxyl at position 1 in β-D-glucose are not involved (Figure 60). Modelling of other carbohydrate guests such as methyl α-D-glucoside, D-mannose and D-galactose predicted that the relatively small cavity might disfavour such guests due to steric interactions, as well reduced numbers of and weaker (i.e. increased intermolecular distance between) hydrogen bonding interactions. These results already suggested some degree of selectivity could be achieved for D-glucose over other carbohydrates.

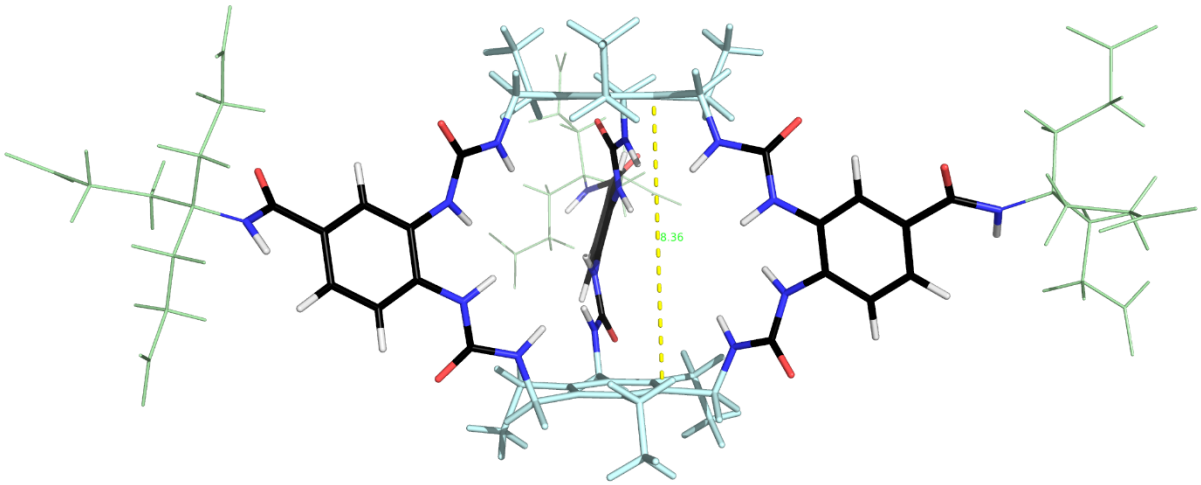
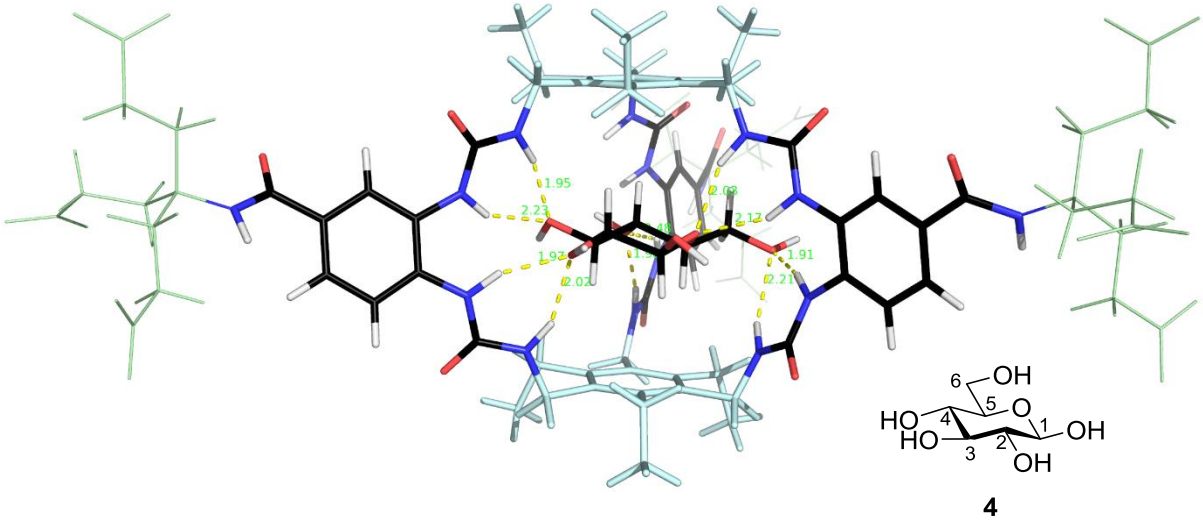


Figure 59 Ground state structure from MCMM of hexaurea **3**. All ureas are predicted to point inwards towards the cavity, with the distance between aromatic surfaces predicted at 8.4 Å. Force field used: OPLS2005 with aqueous GB/SA solvation.

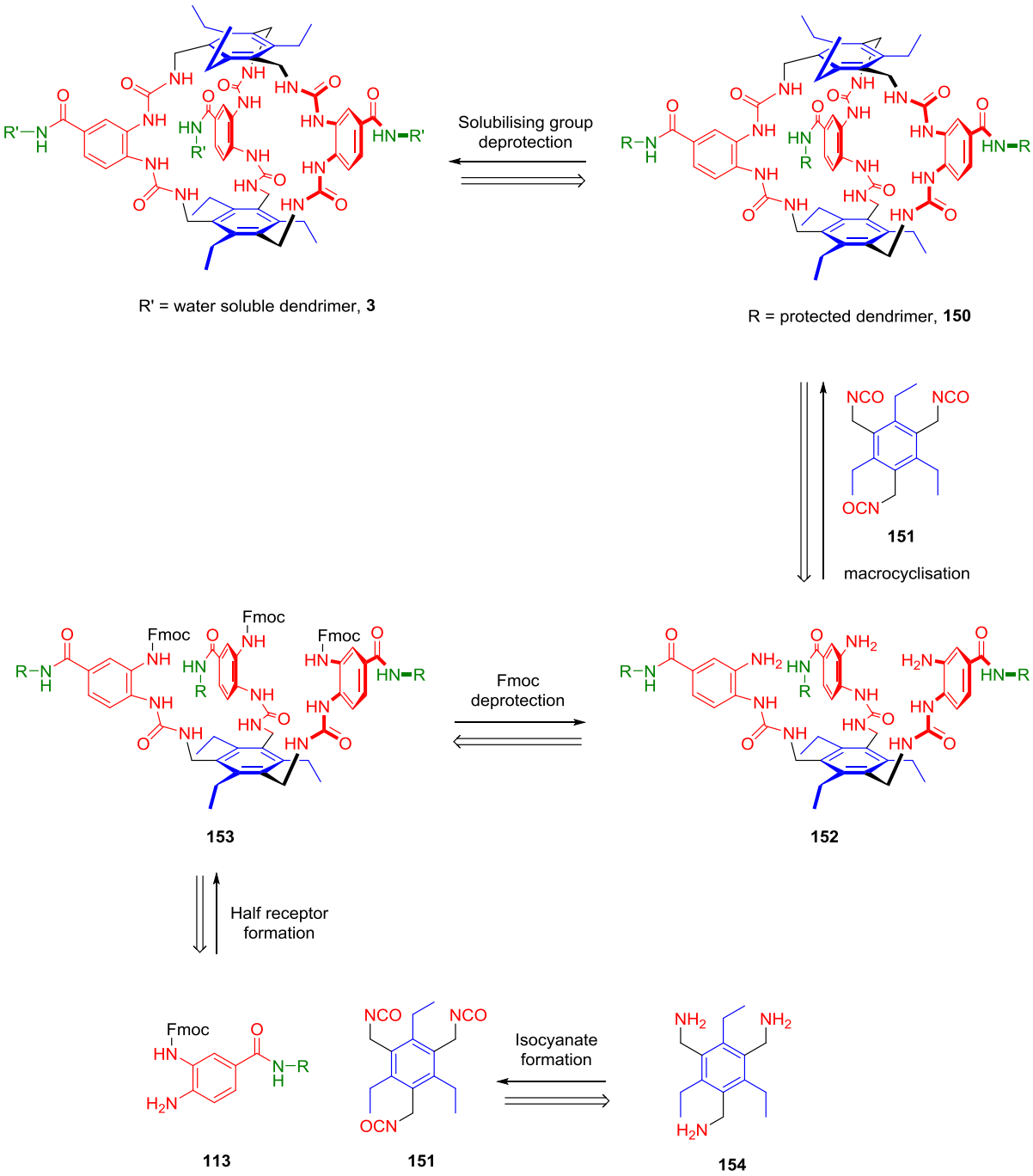


4

Figure 60 Ground state structure from MCMM of hexaurea **3** complexed with β-D-glucose **4**. Ten intermolecular hydrogen bonds are predicted (shown as yellow broken lines), with lengths of 1.95 – 2.48 Å. Only one urea and the hydroxyl at the 1 position on glucose do not take part in hydrogen bonding. Force field used: OPLS2005 with aqueous GB/SA solvation.

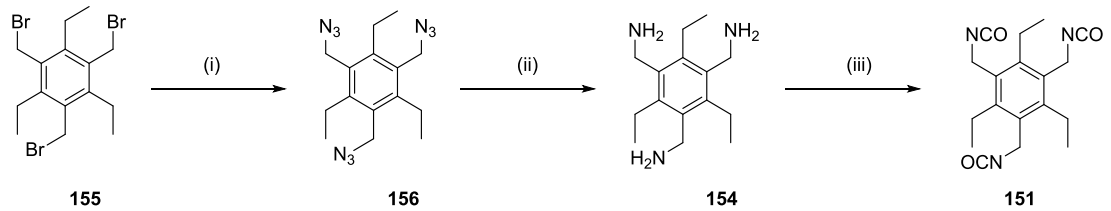
The increased stability of the TEB scaffold when compared with anthracene was also desirable. There was predicted to be a much lower risk of oxidation upon reaction with singlet oxygen, which was expected to enable long term stability in the solid state and even in solution – something of paramount importance for potential practical application of carbohydrate sensing. It was also hypothesised that the single benzene ring of **3** would be less prone to the same acid instability that afflicted the anthracene class of urea receptors (Chapters 2 and 4). The loss of adjacent conjugated aromatic rings was predicted to make cleavage of the urea *via* a benzylic cation much less favourable, and so was hypothesised to be less of an issue for **3**.

Retrosynthetic analysis predicts a similar route as for anthracene receptor **1** (chapter 4), and ultimately leads to two major components: previously synthesised linker **113** and *tris*-isocyanate **151** (Scheme 71). Isocyanate **151** can be synthesised from triamine **154**, and then macrocyclisation *via* the half receptors **153** and **152** will enable access to protected macrocycle **150**, with subsequent deprotection of the solubilising groups to afford water soluble macrocycle **3**.



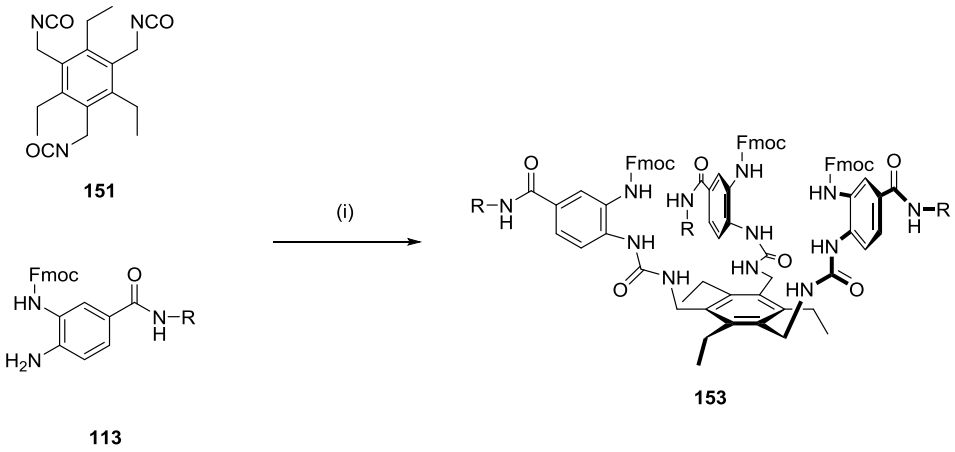
Scheme 71 Retrosynthetic analysis of hexaurea **3**, leads to two key components: previously synthesised linker **113** and *tris*-isocyanate **151**, which was predicted to be synthesised from the equivalent triamine **154**. R = *tert*-butyl protected G2MM solubilising group. R' = water soluble G2MM solubilising group.

5.3 Synthesis and characterisation



Scheme 72 (i) NaN_3, MeCN, reflux, 16h, 92%, (ii) PMe_3, H_2O, THF, rt, 16h 90%, (iii) triphosgene, $NaHCO_3$ (0.1 M aqueous), CH_2Cl_2, rt, 1h 78%.

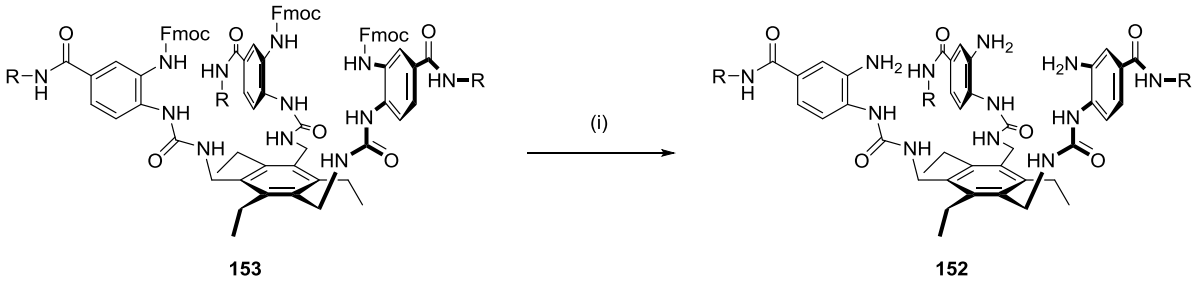
Synthesis of the *tris*-isocyanate **151** began with commercially available *tris*-bromide **155**, which was converted to the *tris*-azide **156** in excellent yield using sodium azide in DMF (Scheme 72). Reduction of **154** to *tris*-amine **154** was achieved also in excellent yield using trimethylphosphine and water in THF. Conversion of **154** to the corresponding isocyanate **151** was achieved using triphosgene in a biphasic mixture of dichloromethane and aqueous sodium hydrogen carbonate, this afforded **151** in good yield. A slight excess of triphosgene was needed to obtain the pure isocyanate with exact stoichiometry resulting in formation of urea by-products – although these by-products could be removed by precipitation and filtration in cold toluene or dichloromethane. The isocyanate group was confirmed by infra-red spectroscopy, where a very intense band attributed to the NCO stretch was observed at 2243 cm^{-1}.



Scheme 73 (i) THF, reflux, 4 days, 56%. R = G2MM solubilising group.

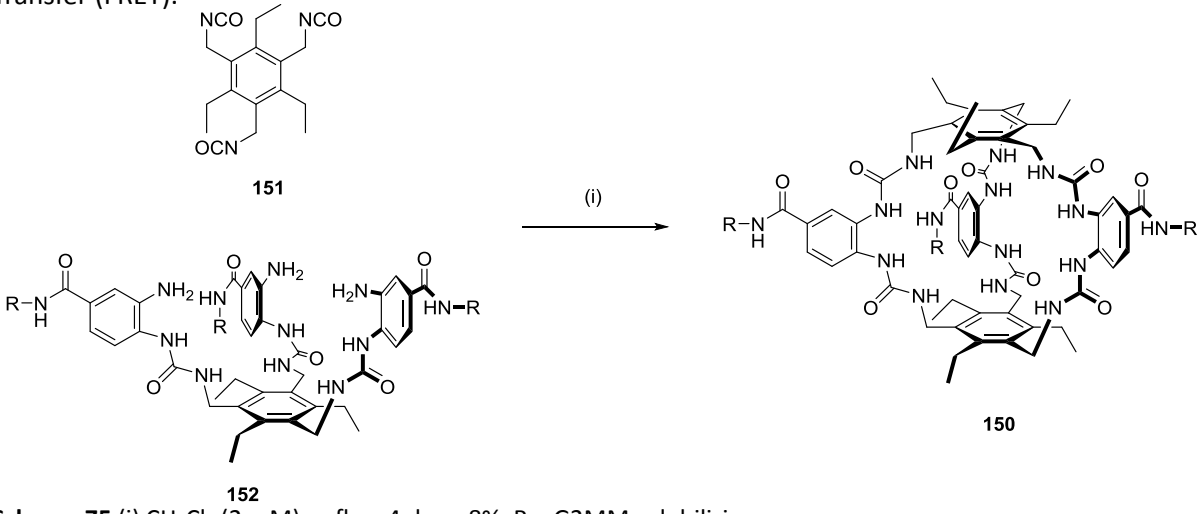
Reaction of linker **113** and isocyanate **151** under reflux in THF afforded protected half receptor **153** in a moderate yield. The reaction proceeded very slowly, taking more than 3 days to consume all isocyanate **151**. This could potentially be due to increased steric bulk near the isocyanate once two linker units had already reacted, making addition of the third bulky amine much slower. Purification on silica using flash column chromatography proved difficult, but separation was achieved using

reverse phase preparative HPLC. The protected half receptor also displayed broadening in the ^{1}H NMR spectrum as seen for anthracene Fmoc-protected half receptors **118** and **128** (see chapter 4). It was speculated that this is due to the same reasons as well: where slow rotating conformers interconvert on the NMR timescale, although this would have to be confirmed with variable temperature NMR studies.



Scheme 74 (i) DBU, CH_2Cl_2, 0 °C, 2h, 92%. R = G2MM solubilising group.

Deprotection of Fmoc half receptor **153** to **152** was achieved in excellent yield using DBU in dichloromethane. Severe line broadening was also observed in the ^{1}H NMR spectrum when using aprotic solvents (such as $CDCl_3$ or CD_2Cl_2), believed to be sign of possible self-association between the urea and aniline groups or intermolecular aggregation. Using CD_3OD as solvent appearing to disrupt this to provide well resolved spectra, suggesting that CD_3OD suppresses the intramolecular interactions observed in the aprotic media. Interestingly, *tris*-amino half receptor **152** was fluorescent when exposed to 365 nm UV light, emitting blue/purple light in solution. This is speculated to be a result of the 'push-pull' conjugation of the amine and carbonyls on the urea spacer units. This emission may have been masked by the inherent fluorescence of anthracene in previously synthesised half receptors (Chapter 4), or even absorbed by the anthracene units *via* a Förster Resonance Energy Transfer (FRET).[94]



Scheme 75 (i) CH_2Cl_2 (2 mM), reflux, 4 days, 8%. R = G2MM solubilising group.

With both the amino half receptor **152** and *tris*-isocyanate **151** in hand, the macrocyclisation to generate hexaurea **150** was attempted (Scheme 75). Rate of reaction initially appeared very slow upon monitoring of the reaction by TLC and analytical HPLC. Additional equivalents of isocyanate were added throughout to promote consumption of half receptor, although whether this consumption indicated cage formation or simply oligomerisation was initially uncertain. Once complete consumption of **150** was observed, purification by reverse phase HPLC was attempted. However, neither the product nor any distinct intermediates were discernible within the chromatogram, as all eluting peaks were very broad with minimal separation (Figure 61). Isolation of these peaks and analysis by ^{1}H NMR suggested all products were very similar, featuring broad resonances with very similar chemical shifts – suggesting these were oligomers of varying lengths.

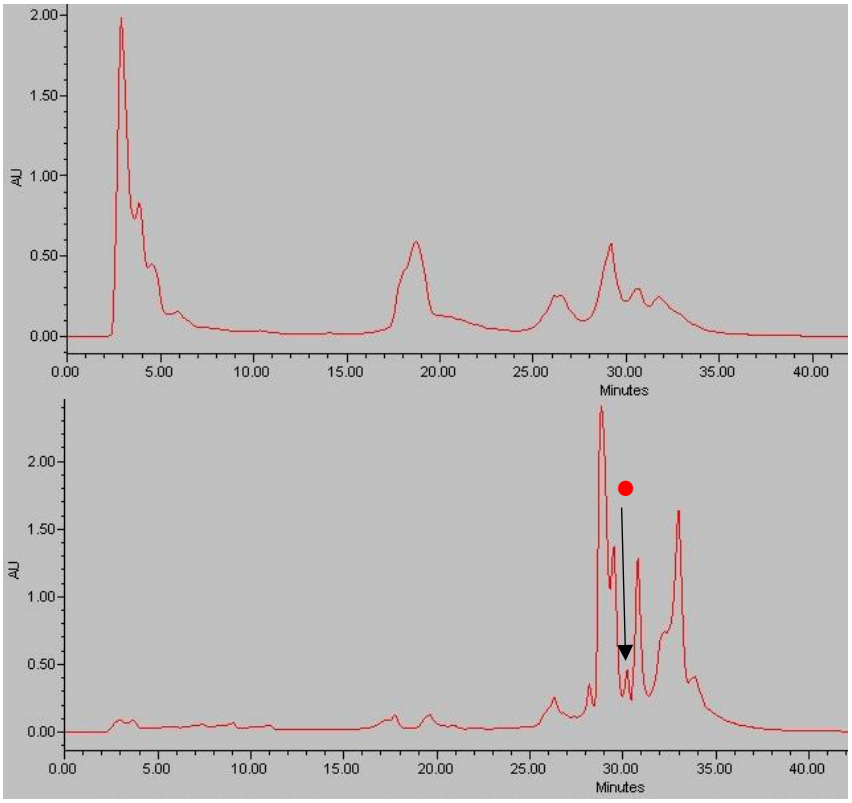
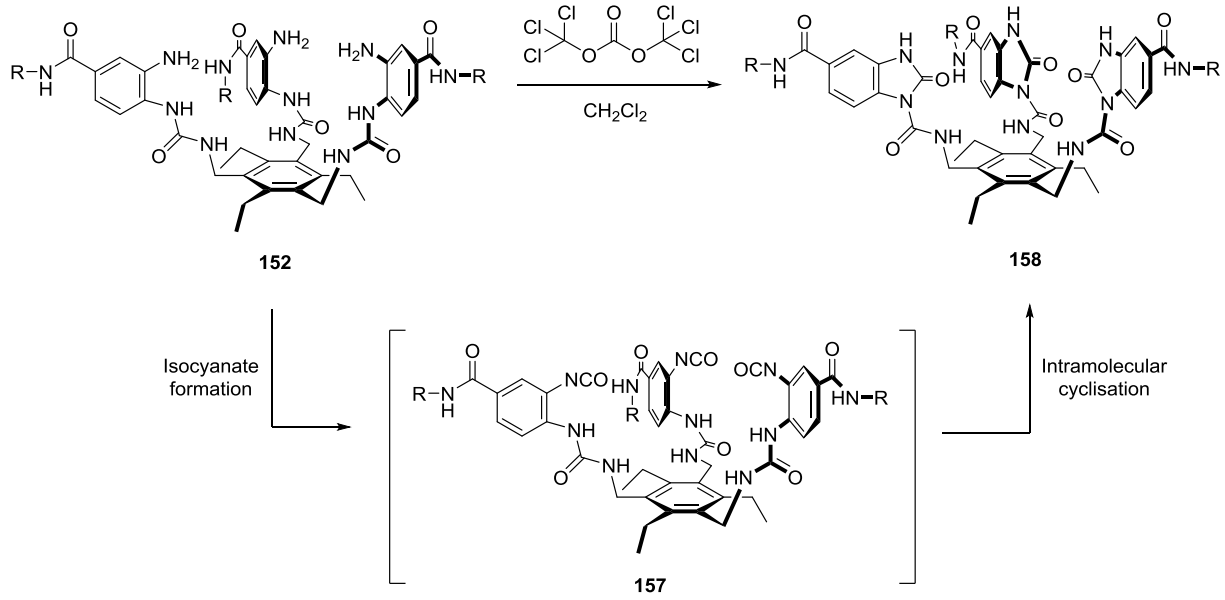


Figure 61 Comparison of HPLC traces for initial macrocyclisation attempts, eluting with acetone/water. The top chromatogram was an early attempt showing the typical broad and poor resolved peaks observed, no macrocycle was isolated. Later optimisation yielded the bottom chromatogram, the peak relating to macrocycle **150** is denoted with •. Isolated yield of **150** was 8%. Detection wavelength of 210 nm was used for both chromatograms. Stationary phase: preparative scale C18, flow rate: 16 mL/min, eluents: acetone/water.

Several reaction conditions were screened in an effort to optimise the macrocyclisation, with desired outcomes including faster reaction rates and cleaner reactions to aid in purification by HPLC. Reaction concentrations were increased from 0.1 mM to 1 mM and then 5 mM to increase reaction rate. Different solvents were also investigated, although it was found that THF, DMF and mixtures like

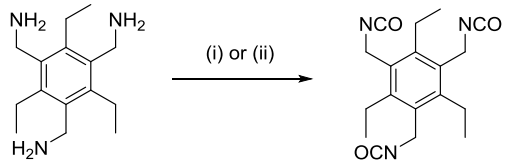
pyridine/dichloromethane gave very convoluted chromatograms. It was also speculated at this time that the *tris*-isocyanate **151** was unstable after isolation and decomposed in the solid state after a few days. In order to circumvent the apparent issues with storing isocyanate **151**, this reagent was subsequently prepared immediately prior to use, as decomposition of **151** to amines or ureas would greatly increase the number of potential products generated. It was eventually found that refluxing in dichloromethane (2 mM reaction concentration) for 4 days, using fresh isocyanate, gave a relatively clean chromatogram with somewhat resolved peaks (Figure 61). Isolation of the peaks and analysis by ^{1}H NMR was non-trivial. Several of the isolated products displayed very broad spectra but a few displayed more defined signals. These were potentially indicative of macrocyclic species. Analysis of these selected fractions by MALDI mass spectrometry revealed that one of them was indeed the desired macrocycle **150**, albeit the fraction with the lowest isolated yield of 8%.



Scheme 76 Conversion of amino half receptor **152** to *tris*-isocyanate **157** by residual triphosgene. Intramolecular cyclisation then affords *tris*-cyclic urea **158**. R = G2MM solubilising group.

Unfortunately, it was found that synthesis of macrocycle **150** could not be replicated. The predominant product isolated was *tris*-cyclic urea **158** (>20% isolated yield, Scheme 76), which featured a ^{1}H NMR spectrum of well resolved peaks with characteristically different chemical shifts for the spacer aromatic protons when compared to macrocycle **150**. The product was ultimately confirmed as **158** by high resolution ESI mass spectrometry, where the $[M+3Na]^{3+}$ was observed as 1731.6936 (calculated value: 1731.6922). This m/z value is also the same as for *tris*-isocyanate **157**, but it was expected that such a reactive species would not persist in the high energy conditions of mass spectrometry. It was confirmed to not be **157**, as IR spectroscopy showed no characteristic stretch for the NCO group. The ^{13}C NMR also revealed two different urea environments. It was theorised that **158** does arise from the

formation of *tris*-isocyanate **157** however, and that **157** was formed by reaction with residual triphosgene left over from synthesis of TEB isocyanate **151**. It was noted during synthesis of TEB isocyanate **151** that the triamine **152** was not completely soluble in dichloromethane, which may lead to incomplete reaction of the amine with triphosgene. This coupled with the slight excess of triphosgene used to drive the reaction may enable some triphosgene to be carried through into the next step. This discovery could also relate to the complex mixture of products observed during purification by HPLC, as partial conversion of half receptor **152** to a mono or bis-isocyanate would increase the number of potential side products formed. Upon reviewing this, it was decided to revisit the synthesis of TEB isocyanate **151** to minimise any leftover triphosgene and enable a cleaner macrocyclisation.



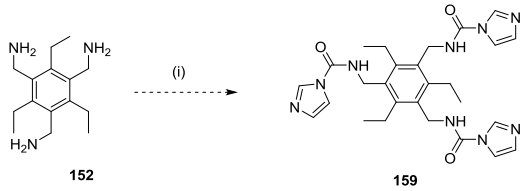
152 **151**

Scheme 77 (i) Triphosgene, EtOAc, $NaHCO_3$ (0.1 M, aqueous), 0 °C, 1h, 32%; (ii) triphosgene, toluene, reflux, 3h, 91%.

The original procedure for the synthesis of **151** was revisited. The solvent was changed from dichloromethane to ethyl acetate, as it was noted that triamine **152** was much more soluble in ethyl acetate (Scheme 77). A stoichiometric amount of triphosgene was also used (one equivalent of triphosgene per triamine **152**). Extraction of the reaction mixture after 1 hour with ethyl acetate afforded very pure **151**, although the yield had been reduced considerably to 32%. The procedure was very consistent, with no isolated yield ever exceeding 32%. It was speculated that the improved solubility of triamine **152** enables it to react with the isocyanates as they are being formed, leading to an increase in urea by-products.

A new procedure was then developed in collaboration with Ziylo Ltd. researchers Mike Orchard and Johnathan Matlock. Triamine **152** was added to an excess of triphosgene in toluene, which was then refluxed for 3 hours (Scheme 77). Removal of the solvent and extraction with cold toluene or dichloromethane then afforded very pure *tris*-isocyanate **151** in excellent yield. All of **151** used henceforth was made using this method. Isocyanate **151** also now appeared much more stable in the solid state than when previously synthesised, which would suggest that unknown impurities from the previous biphasic procedure were responsible for the previously observed instability. The reaction is believed to proceed by initial addition of triamine **152** to triphosgene to give the carbamoyl chloride intermediate, with the excess triphosgene preventing intermolecular reaction between free amines

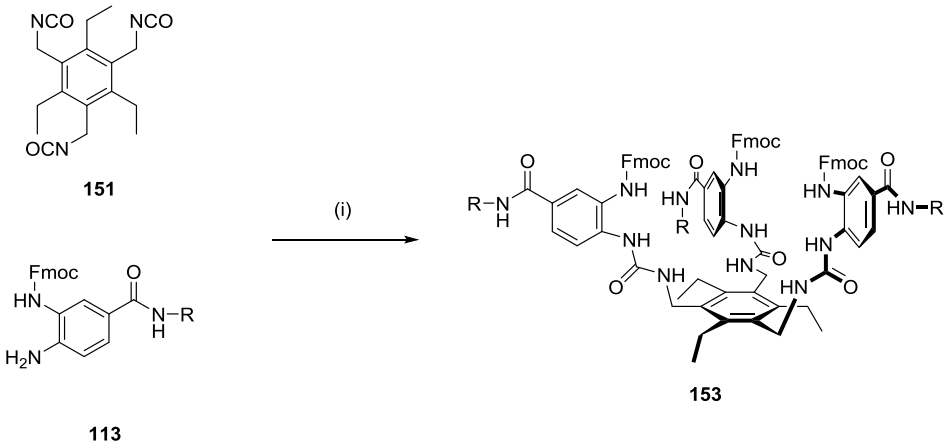
and these intermediates. Upon refluxing, thermal decomposition of the carbamoyl chloride to the isocyanate occurs yielding the product **151**.[95] Excess triphosgene is decomposed to phosgene, which is then evaporated by the high temperatures of the reaction but also by exposure of **151** to high vacuum. Test studies of refluxing triphosgene in toluene revealed that no triphosgene or phosgene remained after 3 hours – suggesting that no phosgene type reagents remained after isolation of **151** using this method.



152 **159**

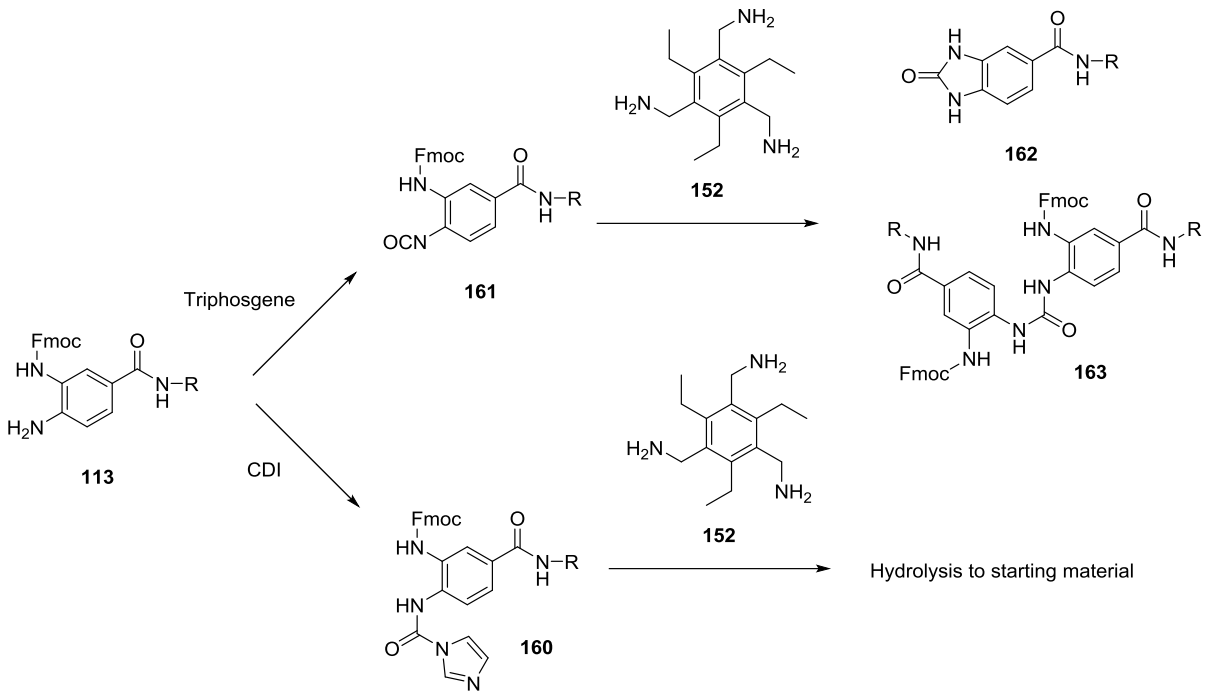
Scheme 78 (i) Carbonyldiimidazole, THF, 45 °C, 2h, **159** not isolated.

In search of a more optimal procedure, conversion of **152** to *tris*-imidazole urea **159** using carbonyldiimidazole (CDI) was also attempted (Scheme 78) to avoid issues related to residual triphosgene. However, shortly after addition of CDI to the reaction a white precipitate was observed – initially thought to be an indication of imidazolium formation. TLC of the reaction mixture implied more than one product had formed. Filtration of the reaction and analysis of the solid by ^{1}H NMR revealed the precipitate to be a mixture of mono and bis-imidazole ureas, as well as imidazole. The filtrate was found to contain predominantly imidazole, although mass recovery was low for the filtrate. It was then hypothesised that the mono and bis-urea intermediates are insoluble in the reaction mixture and cannot progress fully to the product **159**. No further investigations were attempted as the procedure to give pure TEB isocyanate **151** had been optimised by this point (Scheme 77).



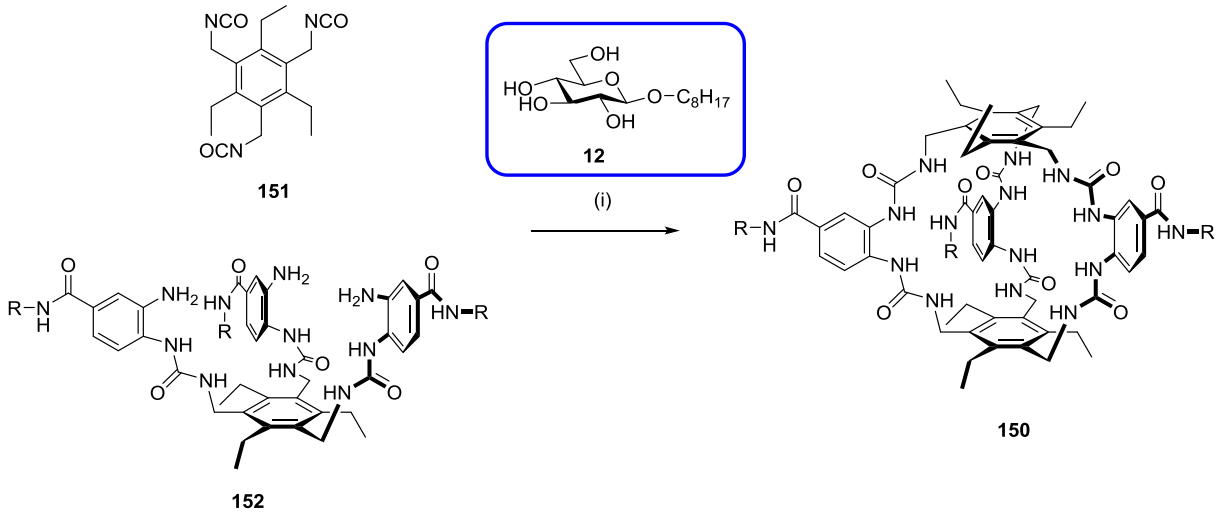
Scheme 79 (i) pyridine, CH_2Cl_2, reflux, 16h, 84%. R = G2MM solubilising group.

With TEB isocyanate **151** synthesised using the new procedure, previous reactions to synthesise Fmoc protected half receptor **153** were no longer reproducible. No reaction was observed when refluxing linker **113** and **151** in both THF and dichloromethane. It was deduced from this that trace impurities from the previously used isocyanate synthesis (e.g. inorganic bases such as carbonate) were potentially catalysing the reaction, and this impurity was now absent using the new isocyanate procedure. Several additives were then screened to replicate this previous activity. It was found that no product was formed upon addition of organic bases such as DIPEA, although consumption of isocyanate was observed – this was most likely due to hydrolysis to the amine and then urea formation. Addition of the nucleophilic catalyst DMAP appeared to promote a relatively rapid reaction, with new products clearly appearing by TLC. However, also evident was Fmoc deprotection of both products and starting materials. This yielded a complex mixture of products by TLC, with attempted purification by column chromatography on silica proving non-trivial. Fortunately, pyridine was found to both accelerate the rate of reaction and keep the Fmoc protecting groups intact. This afforded a very clean reaction that gave protected half receptor **153** in excellent yield after column chromatography on silica. It was found that an eluent of 1:1 EtOAc/dichloromethane allowed elution of excess linker **113** and flushing of the column with 5% methanol in dichloromethane eluted pure **153**. Excellent separation was also achieved by reverse phase flash chromatography using a Biotage Isolera One, with a C18 stationery phase and eluting with acetone/water. Deprotection of the Fmoc groups using DBU to give amine half receptor **152**, as discussed before, was trivial and still occurred in excellent yields (Scheme 74).



Scheme 80 Attempted conversion of linker **113** to an electrophile and reaction with triamine **152** did not yield half receptor **153**. Only hydrolysis of the imidazole urea **160** was observed. Whereas isocyanate **1161** yielded dimerised product **163** and cyclic urea **162**.

In conjunction with screening additives to promote reaction to the half receptor **153** (Scheme 79), the reverse approach was also attempted – where linker **113** would be converted to an electrophile and subsequent reaction with triamine **152** would yield half receptor **153**. Imidazole urea **160** was prepared *in situ* from **113** using CDI, with formation inferred by TLC. Upon addition of amine **152** however, no product was observed by TLC analysis, with mainly starting material recovered upon purification. Hydrolysis of the imidazole urea must have occurred, with the triamine **152** introducing water to reaction as it would be predicted to be hygroscopic. A very different outcome was observed for reaction of isocyanate **161**, which was converted from **113** using triphosgene. Analysis by TLC and IR spectroscopy indicated conversion to the isocyanate. After addition of triamine **152** to **161**, two new products were clearly visible within a matter of hours. Isolation of these products and analysis by NMR and mass spectrometry, suggested the structures of cyclic urea **162** and dimerised product **163**. Hydrolysis of the isocyanate to the aniline and then reaction with another isocyanate would yield dimerised urea **163**. Intramolecular cyclisation of the Fmoc-carbamate onto the isocyanate would yield a cyclic urea (see Scheme 5, Chapter 2 for a similar example). It is postulated that the Fmoc group becomes more labile in this intermediate and is then lost, potentially by deprotection from triamine **152**, as no characteristic indications of the Fmoc group were visible in the product by NMR. No further investigation of this route was then attempted, as synthesis of **153** by another route had been optimised by this point.



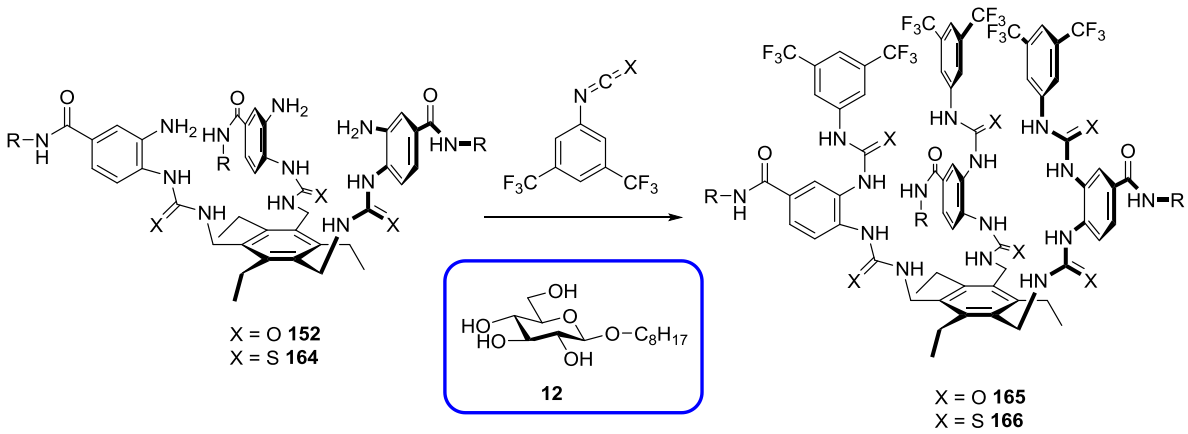
Scheme 81 (i) octyl β-D-glucoside **12**, DMAP, CH_2Cl_2 (0.6 mM reaction concentration), reflux, 2 days, 48%. Reaction yield was greatly improved upon addition of octyl β-D-glucoside **12**, which templates formation of macrocycle **150**. R = G2MM solubilising group.

TEB isocyanate **151** synthesised using the new procedure also proved unreactive when attempting to form macrocycle **150** from **151** and **152**. With lessons learned from the synthesis of **153** and the critical discovery of using pyridine to promote reaction, several additives were screened for the macrocyclisation to give **150**. As for Fmoc half receptor **153**, organic bases such as DIPEA and even DBU showed no reaction – TLC and analytical HPLC showed only half receptor **152**, with no products formed. Addition of nucleophilic catalysts such as DMAP and pyridine however rapidly increased the rate of reaction, with DMAP providing the greatest increase of rate. No reaction control was evident however, with very complex chromatograms being produced (Figure 62).

When octyl β-D-glucoside **12** was added (with DMAP present) a dramatic improvement in reaction control was observed, with a much cleaner reaction observed by TLC and analytical HPLC. A peak of large intensity at the correct retention time of **150** was observed. Indeed, isolation of this fraction by reverse phase HPLC (acetone/water eluent) revealed it to be macrocycle **150** in much improved yield (~30%). Furthermore, this result proved highly reproducible with isolated yields typically in the range of 30-40%. These results suggested that octyl glucoside is templating the formation of the macrocycle **150**. Fortuitously, the octyl glucoside **12** did not appear to interfere with HPLC purification either, eluting very early along with the solvent front and DMAP, and was found to not be complexed to the receptor upon isolation. No evidence that DMAP facilitated reaction of octyl β-D-glucoside **12** and TEB isocyanate **151** to give a carbamate was found either through isolated products or control studies with NMR spectroscopy.

NMR binding studies between half receptor **152** and octyl glucoside **12** gave a K_a of ~300 M^{-1} in CD_2Cl_2 (1:1 binding model). This association constant is relatively weak when compared to similar receptor

architectures reported by the groups of Roelens and Mazik – these typically feature association constants in the range of 10^3-10^5 M^{-1} in $CDCl_3$.[23,25] The spectra of **152** were much broader in CD_2Cl_2 compared with CD_3OD, which may imply some degree of intramolecular interaction between the urea groups and the anilines. Dilution of **152** appeared to have no effect on the spectra which disfavours the theory that the broad spectra are due to intermolecular aggregation. It is speculated that the lower binding to octyl glucoside **12** is due to the fact that binding of **12** would require disruption of these intramolecular hydrogen bonds. This would lower the net thermodynamic gain (i.e. ΔG of binding) upon binding of octyl glucoside **12**, thus lowering the association constant. Indeed, such a hypothesis was further reinforced upon synthesis of acyclic receptors **165** and **166** by visiting student Eléa Wallisky (Scheme 82). These receptors exhibited even lower binding constants (than **152**) to octyl glucoside **12** in $CDCl_3$, and it was rationalised that conversion of the aniline groups to ureas strengthened the aforementioned intramolecular hydrogen bonds which inhibit intermolecular binding of **12**.



Scheme 82 Synthesis of acyclic hexaurea receptors **165** and **166** from **152** and **164**, which was carried out by visiting student Eléa Wallisky. Both receptors show very weak binding to octyl β-D-glucoside **12** in $CDCl_3$ (K_a <100 M^{-1}). R = G2MM solubilising group.

An NMR binding study of macrocycle **150** and octyl β-D-glucoside **12**, in CD_2Cl_2, however implied a much higher association constant. The receptor appeared to saturate around roughly one equivalent of glucoside and gave a very convoluted NMR spectrum, presumably due to the asymmetry of the glucoside and receptor complex. No binding constant was able to be determined due to the complex spectra and rate of saturation, but it is speculated to be much higher than that of half receptor **152** although this will need to be quantified in future work. This information lends to the theory that the association constant of octyl glucoside **12** to reaction intermediates increases as formation of the macrocycle proceeds, with each urea linkage formed strengthening the interactions between the intermediate and the glucoside. Once one urea linkage has been formed between half receptor **152**

and isocyanate **151**, the octyl glucoside **12** binds to the new intermediate and then preorganises the conformation towards macrocycle formation. The low binding affinity of octyl glucoside to half receptor **151** might allow unwanted partial oligomerisation to occur however, as there may be no preorganisation before binding of the glucoside to the macrocyclic intermediates. Some evidence of this could be the isolation of dimerised product **167** (Figure 63).

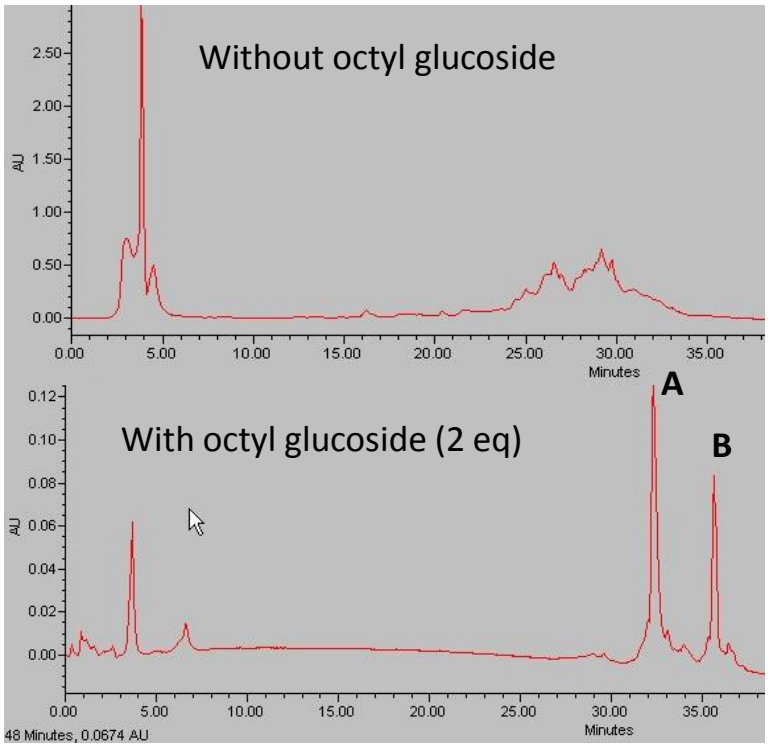
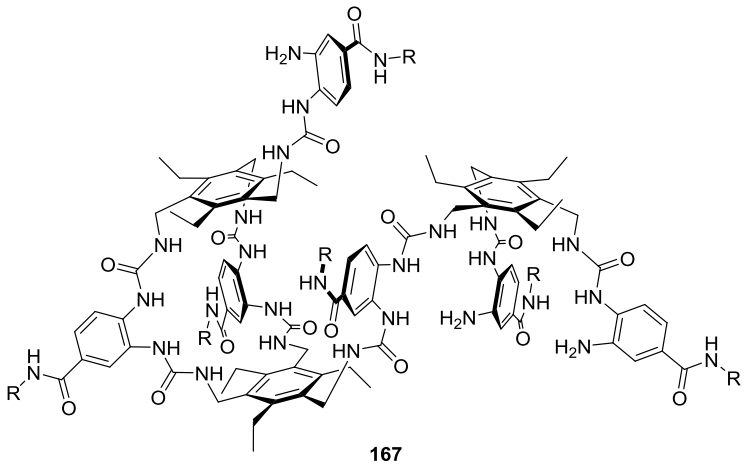


Figure 62 HPLC chromatograms showing the effect of templating the reaction with octyl β-D-glucoside **12** (detection wavelength used is 210 nm). The top chromatogram is a typical trace from initial screening of the macrocyclisation using pyridine or DMAP to accelerate reaction rate but without control of the products formed. None of macrocycle **150** was isolated and very complex mixture of products is observed. The bottom chromatogram is typical of the optimised reaction when octyl β-D-glucoside **12** is present and shows two main products: receptor **150** (A) and a dimerised product **167** (B, see Figure 63 below). Isolated yield of **150** is 48%. Stationary phase: preparative scale C18, flow rate: 16 mL/min, eluents: acetone/water.



167

Figure 63 Hypothesised dimerised product **167** (B, Figure 62) isolated during purification of macrocycle **103**. It consists of 2 equivalents of half receptor **152** and 1 equivalent of isocyanate **151.** Species with similar exact mass observed in mass spectrometry.

The macrocyclisation was then further improved by decrease of reaction concentration to 0.5-0.6 mM and a slow addition of TEB isocyanate **151** *via* syringe pump. It was also found that co-evaporation of half receptor **152**, DMAP and octyl β-D-glucoside with dry toluene prior to reaction provided an even cleaner reaction – possibly due to the removal of water from the more hygroscopic components (such as octyl β-D-glucoside). These modifications gave a final improved yield of ~50%, which again was very reproducible, even on larger scales (>0.5 g).

With macrocycle **150** now routinely accessible, full characterisation of the macrocycle was possible. The ^{1}H and ^{13}C NMR spectra featured similarities when compared with anthracene macrocycle **120** (chapter 4). Chemical shifts and general arrangement of the signals related to the spacer aromatics were largely consistent between the two macrocycles. Also clearly evident by NMR were the two TEB surfaces, which displayed distinct signals in both ^{1}H and ^{13}C NMR (Figure 64). The macrocycle was also confirmed by high resolution ESI mass spectrometry, by observation of a $[M+3Na]^{3+}$ species.

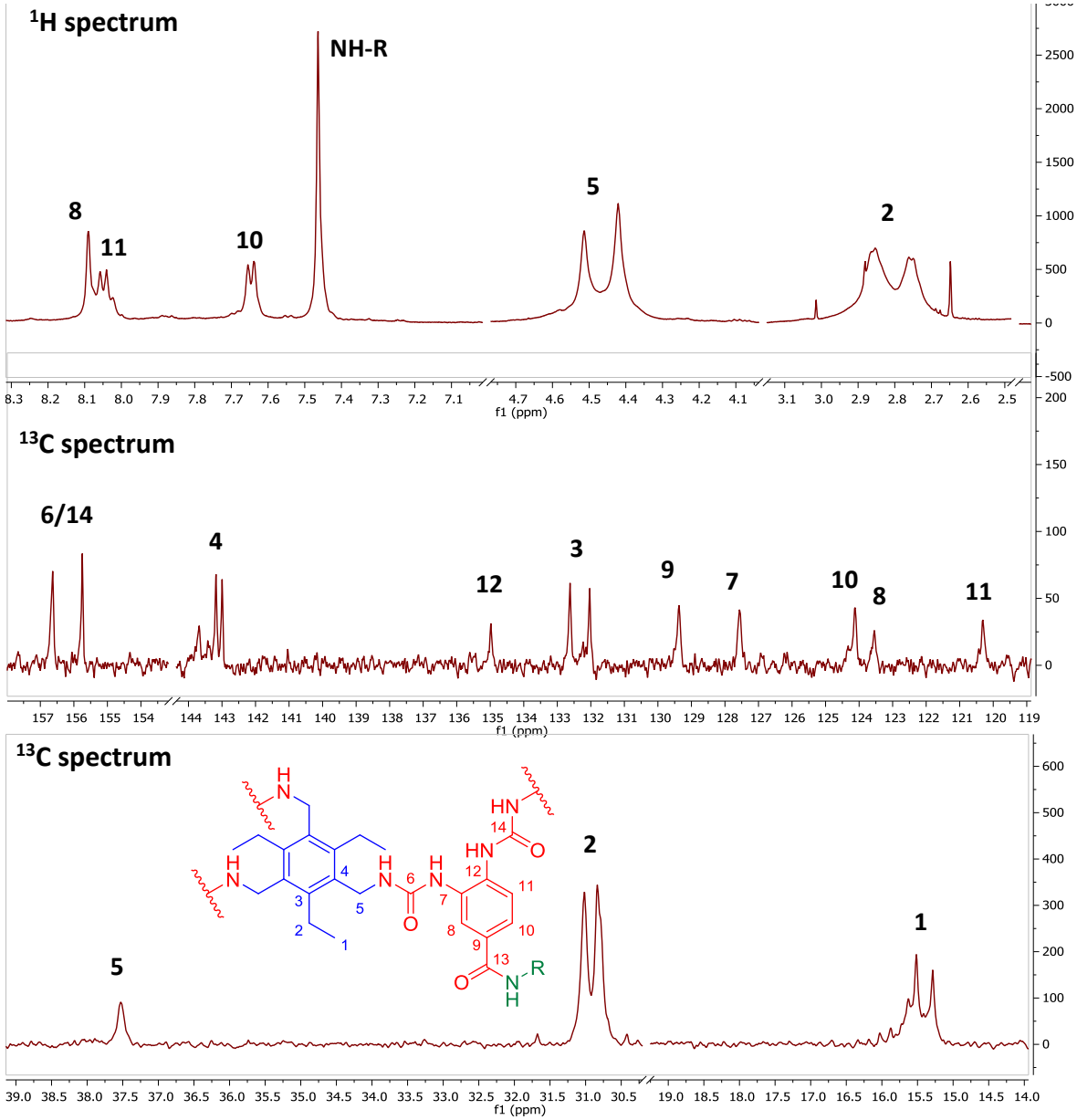
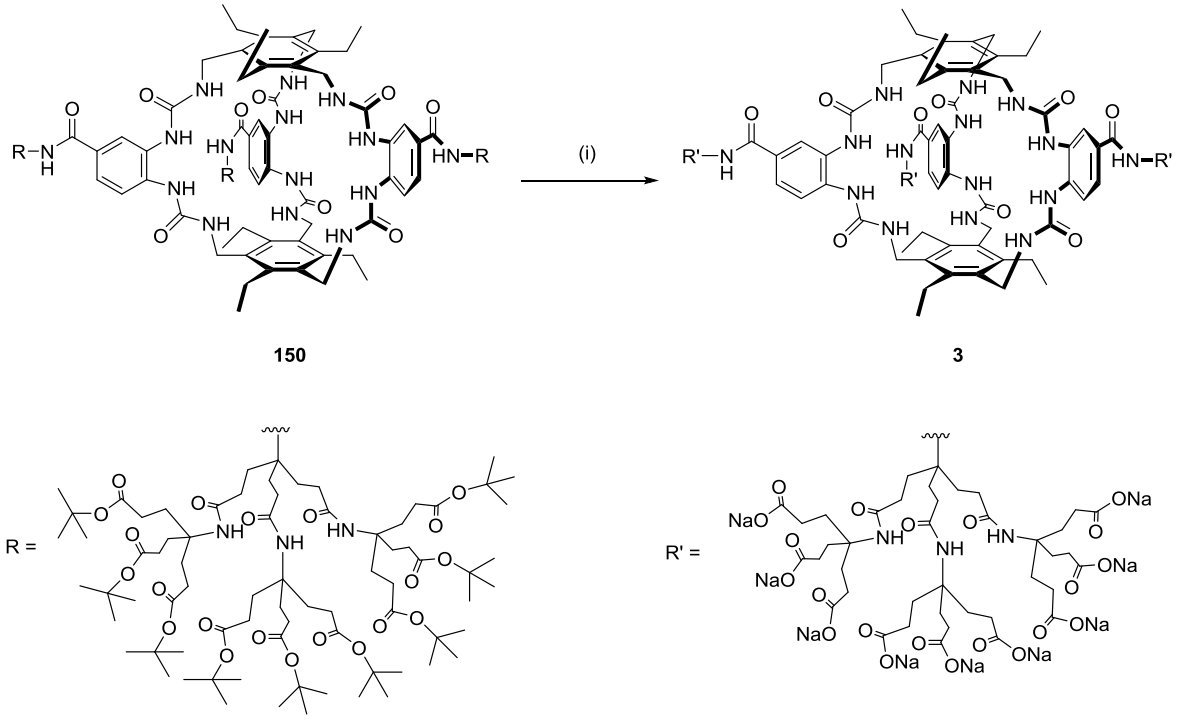


Figure 64 Partial ^{1}H (top) and ^{13}C (middle and bottom) NMR spectra (500 MHz, CD_3OD, 298 K) of macrocycle **150**, demonstrating the inequivalence of the TEB surfaces due to the asymmetric spacer unit.

Deprotection of the *tert*-butyl esters on the solubilising groups with acid was required to afford water-soluble macrocycle **3**. Given the immense amount of work investigating the stability of previous anthracene urea receptors to acid (see chapters 2 and 4), a more cautious approach was taken initially to the acidic deprotection of **150**. However, the first attempt at finding suitable deprotection conditions, which consisted of 20% TFA (v/v) in dichloromethane for 16 hours, did not appear to decompose the macrocycle (Scheme 83). Removal of the TFA and solvent, and then analysis by ^{1}H NMR in CD_3OD revealed that the *tert*-butyls had all been removed but the macrocyclic structure appeared to still be intact, as the spectra displayed a similar composition of chemical shifts and peaks as protected macrocycle **150**. Neutralisation of this product to pH 7 with aqueous sodium hydroxide was

meant to afford polycarboxylate macrocycle **3** and thus water solubility. Indeed, during neutralisation of the reaction product, an insoluble suspension of the polycarboxylic acid (pH ~4) became a solution of polycarboxylate **3** at pH 7. Removal of the solvent and analysis by ^{1}H NMR in D_2O afforded a spectrum consistent with the symmetrical macrocycle.



Scheme 83 (i) TFA (20% v/v), CH_2Cl_2, rt, 16h, 89%. R = *tert*-butyl protected G2MM solubilising group. R' = polycarboxylate G2MM solubilising group.

It was noted however, that small broad peaks not associated with macrocycle **3** were present in the ^{1}H NMR – most notably in the aromatic and urea methylene regions (Figure 65). It was theorised that these impurities may arise from a small amount of decomposition of protected macrocycle **150** during the acidic deprotection, given that the equivalent anthracene macrocycles would completely decompose under the same conditions. Analysis of the product (at pH 4) by reverse phase HPLC using a methanol/water (buffered with 0.1% TFA) eluent, did reveal a major product (i.e. receptor **3**) but also a small quantity of impurity (<10% by integration of the chromatogram). Purification of the receptor using preparative reverse phase HPLC with this same eluent did afford pure receptor **3** (Figure 65), and thus this purification was routinely carried out upon the crude deprotected macrocycle before neutralisation to pH 7. Occasionally, minor esterification of the solubilising chains to the methyl ester would occur in the methanol/water (buffered with 0.1% TFA) eluent. Use of methanol as an eluent was necessary, as poor solubility in acetonitrile was observed and no retention was achieved using acetone. It was determined by integration in the ^{1}H NMR that approximately <30% of the receptors featured one methyl ester. Fortunately, stirring **3** in aqueous base (0.5 M NaOH) removed all traces of

methyl esters. Test studies by NMR also revealed receptor **3** to be stable to 0.5M NaOH (aq) for at least 1 day. The purity of the receptor was then determined by ^{1}H NMR in D_2O, whereby a known concentration of receptor was compared to 1 equivalent of DMF which was added as an internal standard. A long relaxation delay (>45 seconds) was used during acquisition and comparison of the integrals for both receptor and DMF gave a ratio of the two species, and thus an indication of purity (often >80% pure). The impurities not visible by ^{1}H NMR are assumed to be various salts (e.g. sodium trifluoroacetate) that persist through the neutralisation procedure. Accurate determination of purity and thus receptor concentration is critical when attempting binding studies, as the wrong concentration can lead to poor fitting of data with large errors but more importantly can give misleading results.

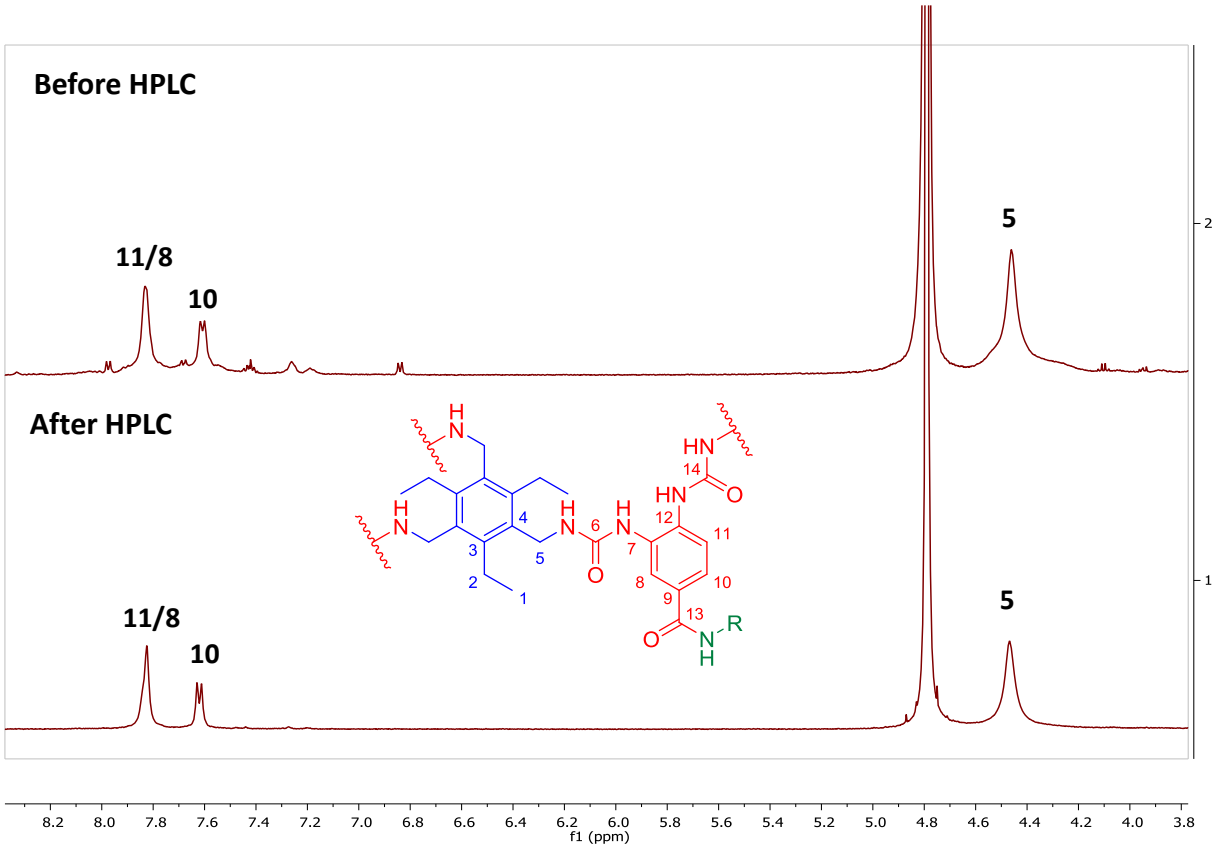


Figure 65 Comparison of ^{1}H NMR spectra (500 MHz, D_2O, pH 7.4, 298 K) of receptor **3** before (top) and after purification (bottom) by reverse phase HPLC. Only the aromatic and urea methylene regions are shown for clarity.

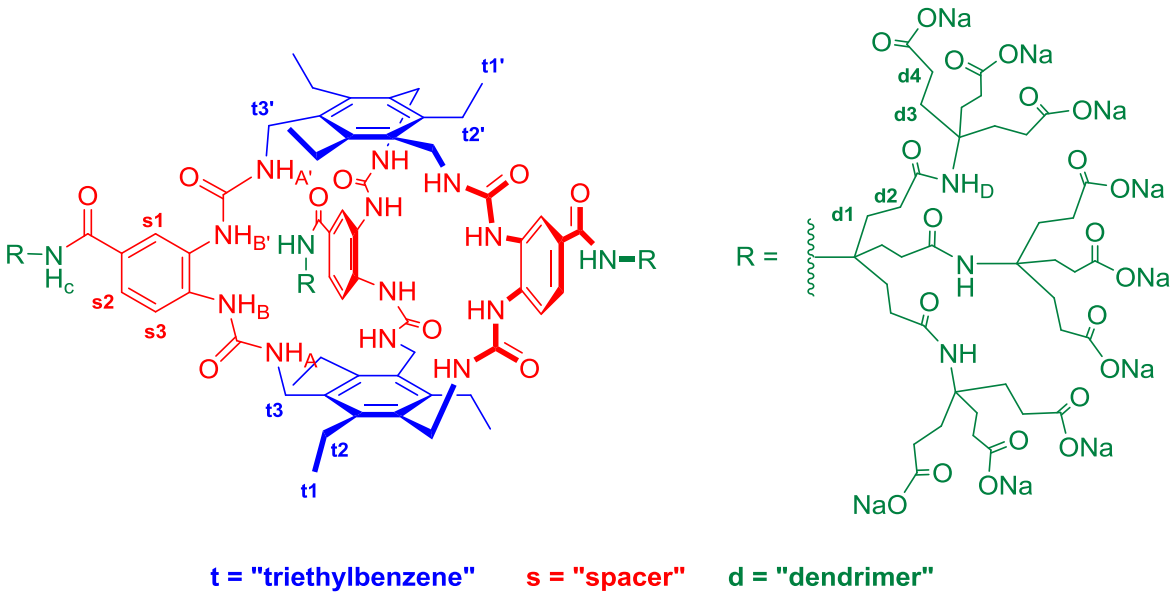


Figure 66 Numbering system used for the discussion of the conformational and binding properties of receptor **3**. Triethylbenzene proton designations (blue) are prefixed with "t" (i.e. t1, t1' etc.), spacer CH protons (red) are prefixed with "s" and protons from the dendritic solubilising groups (green) are prefixed with "d". The urea and amide protons are also denoted accordingly.

Now that receptor **3** could be reproducibly synthesised and purified, full characterisation was achieved using NMR. The numbering system used for discussing characterisation and conformational experiments for receptor **3** is outlined in Figure 66. The ^{1}H NMR was obtained in both D_2O and 9:1 H_2O/D_2O - so that the urea and amide NHs would be visible and reduce exchange with deuterium from the D_2O solvent (Figure 67). The asymmetry of the spacer units and the effect this has on the TEB surfaces could be most clearly seen for the CH_2 (t2) and CH_3 (t1) protons, with both featuring a small broader peak adjacent to the main peak (denoted as t1' and t2'). These major and minor peaks were further characterised by 2D NMR ^{1}H-^{13}C HSQC, where they correlated to carbon atoms with very similar chemical shifts (Figure 69). The 1D ^{13}C spectrum corroborated this with all carbons relating to the TEB surfaces featuring two carbon signals of similar chemical shift (Figure 68) – as seen for protected macrocycle **150**. Furthermore, ^{1}H-^{1}H COSY spectra also revealed that the t1-CH_3 correlated to t2-CH_3, and that t1'-CH_3 correlated to t2'-CH_2, which would infer that the major and minor peaks are indeed from different TEB surfaces (Figure 69).

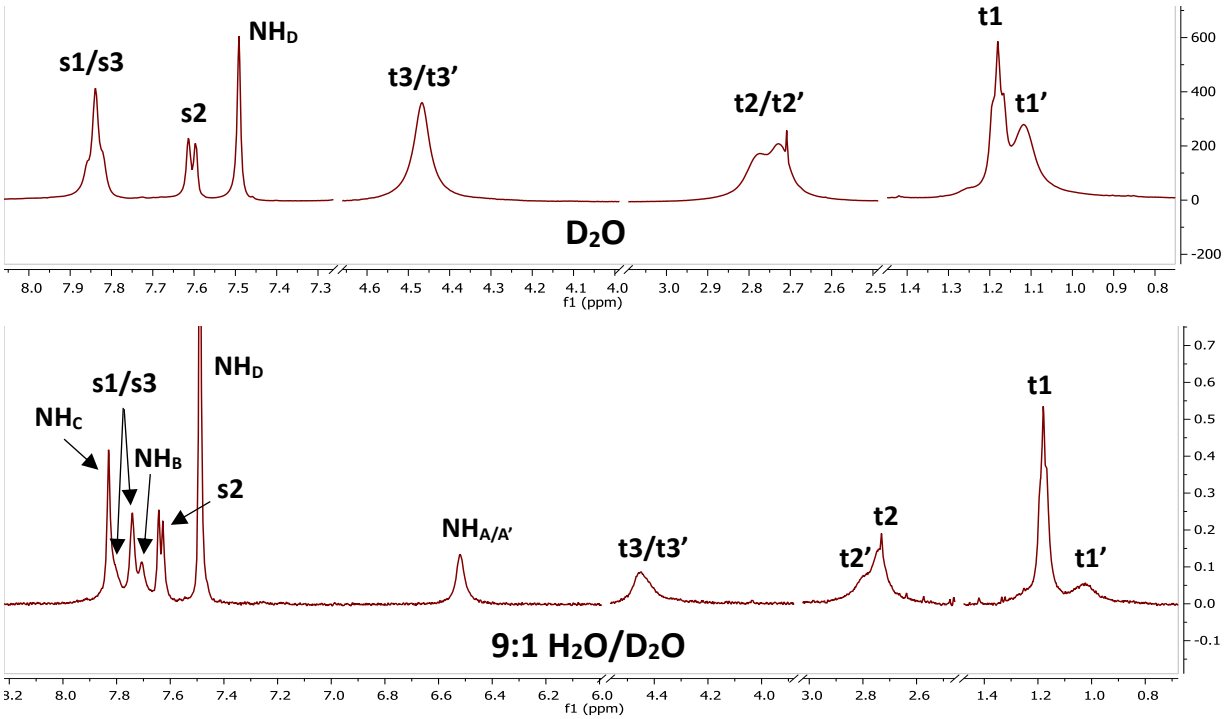


Figure 67 Comparison of partial ^{1}H NMR spectra (600 MHz, pH 7.4, 298 K) of receptor **3** in D_2O (top) and 9:1 H_2O/D_2O (bottom), where the added water reduces deuterium exchange and allows visualisation of the NH protons. Protons related to the solubilising groups and empty space omitted for clarity.

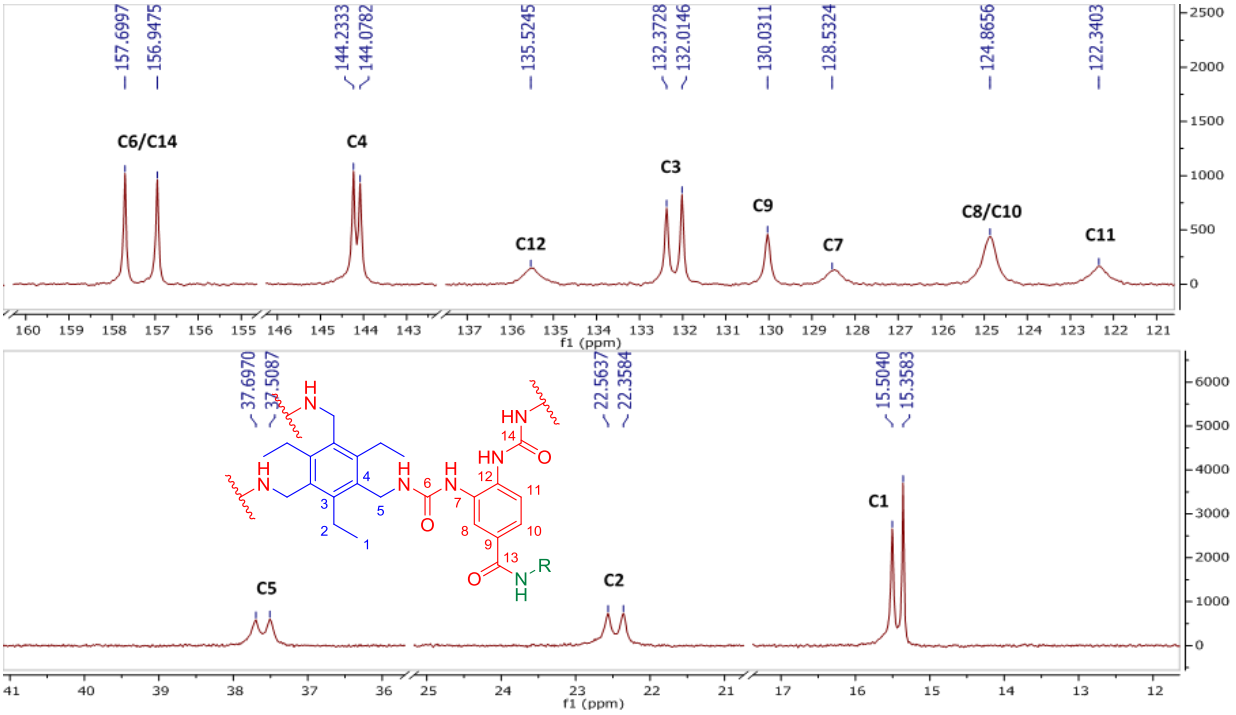


Figure 68 Partial ^{13}C NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) of receptor **3**. The non-equivalence of the triethylbenzene units is reflected in the pairs of signals observed for C1-C5 – as for macrocycle **150**.

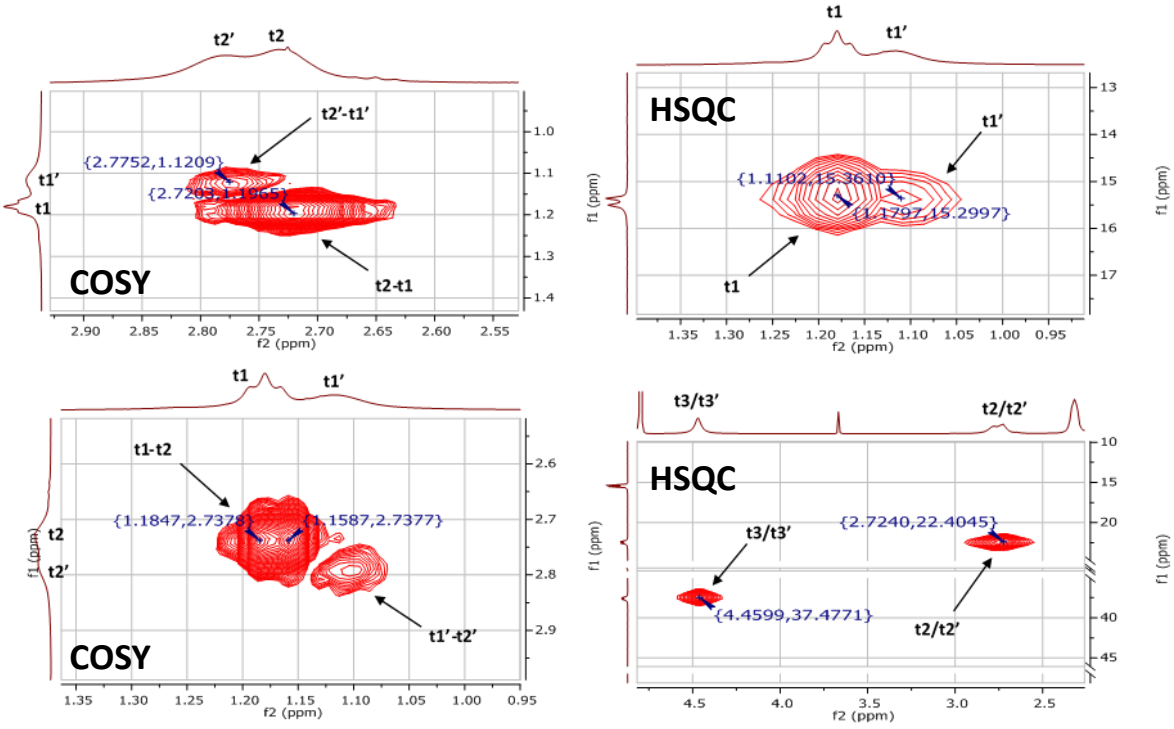


Figure 69 Partial COSY (^{1}H-^{1}H correlation, left) of receptor **3** (600 MHz, D_2O, pH 7.4, 298 K), showing the correlations between major (t1-t2) and minor (t1'-t2') ethyl signals. Partial HSQC (^{1}H-^{13}C correlation, one bond, right) of **3** in D_2O displays that both major and minor ethyl peaks (i.e. t1 and t1') correlate to very similar carbon signals. These data would suggest that the major peaks (i.e. t1 and t2) are from the same TEB surface, and that t1' and t2' originate from the other TEB surface.

Receptor conformation was then probed using 2D NOESY (^{1}H-^{1}H correlation, through space) spectra in 9:1 H_2O/D_2O – so that correlations between the NH protons and other parts of the molecule could be observed. Molecular modelling of the free receptor (Figure 59) predicted a conformation where the urea NH protons were directed inwards towards the centre of the cavity. Analysis of the NOESY spectra appeared to confirm that this conformation was indeed the predominant form of receptor **3**. If rotation of the urea bonds were to occur, then aromatic ureas $NH_{B/B'}$ would be directed towards spacer protons s3 and s1 (Figure 70). However, no such correlation was observed in the NOESY spectrum. The only visible correlations to the aromatic ureas $NH_{B/B'}$ were to the ethyl protons (t1/t1' and t2/t2'), the urea methylenes (t3/t3'), and to the adjacent urea $NH_{A/A'}$. This may suggest that steric interactions between the urea protons and the ethyl groups may prevent rotation and force the ureas to face inwards towards the cavity.

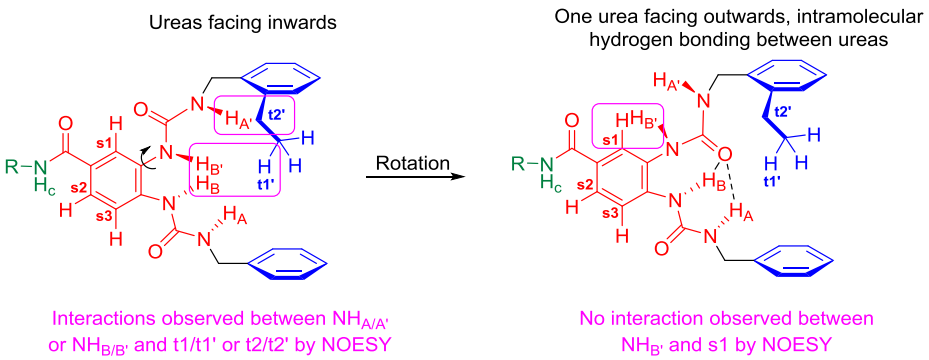


Figure 70 Theoretical conformations of receptor **3** in water. The pink boxes represent possible NOE interactions between protons. The NOESY data infers the conformation on the left, where the ureas are pointing inwards, as through space interactions are observed for the urea and ethyl protons. Evidence for the conformation on the right was not observed, as no interaction was seen for urea $NH_{B/B'}$ and spacer proton s1.

The NOESY spectra also confirmed the aforementioned characterisation of the two distinct TEB ethyl groups, where correlations were only seen between t1 and t2, and between t1’ and t2’ for example. Determination of where each TEB surface (i.e. surface with protons t1 or t1’) is located relative to the asymmetric spacer unit was also achieved – with protons t1’ and t2’ only showing correlations to spacer protons s1. The broad peaks associated with protons t1’ and t2’ could therefore be due to the increased steric bulk on the side of the spacer unit where the dendrimer is attached *via* the amide, as impeded rotation could broaden NMR signals. Conversely, surface protons t1 and t2 only appeared to interact with spacer protons s2 and s3, which would imply t1 and t2 are on the same side as protons s2 and s3. The reduced steric bulk at this part of molecule corroborates with the sharper peaks observed for ethyl group (t1 and t2). These observations are consistent with the structure and numbering system for **3** outlined in Figure 66.

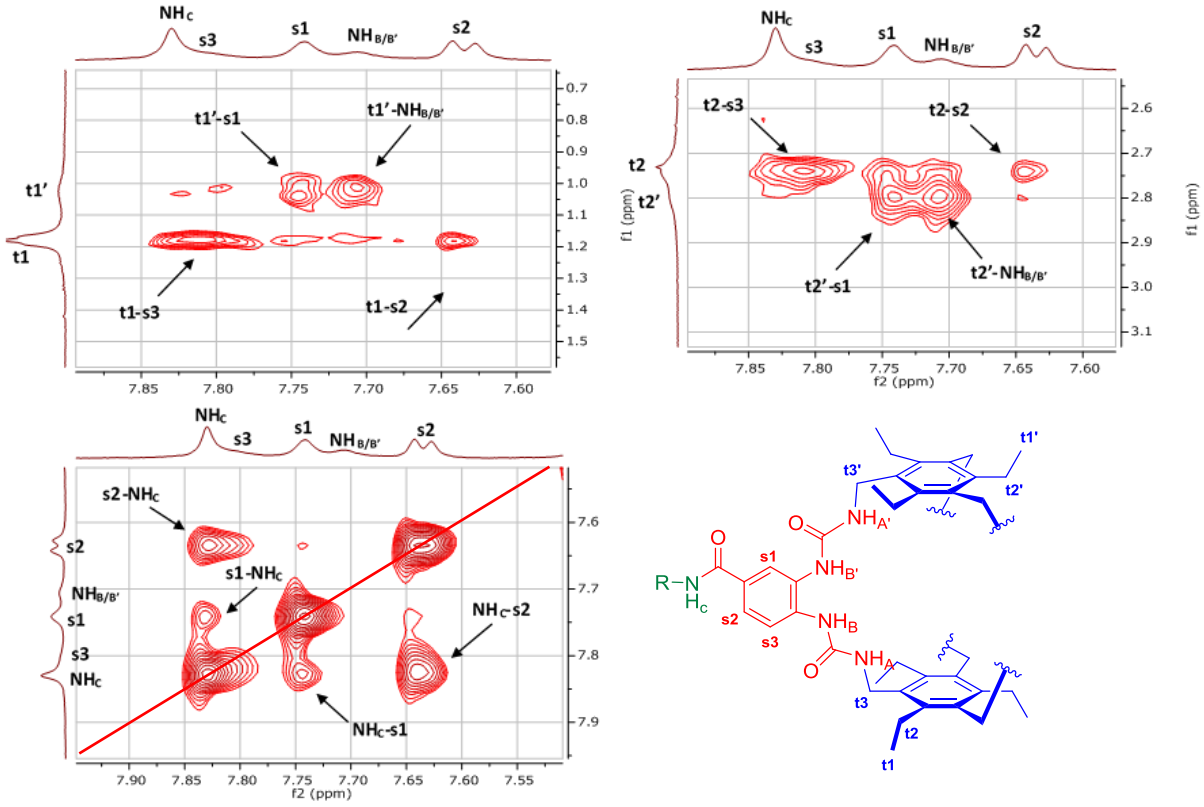


Figure 71 Partial NOESY (^{1}H-^{1}H correlation, through space) spectra for receptor **3** (600 MHz, pH 7.4, 298 K) in 9:1 H_2O/D_2O. The top two spectra confirm the assignment of the receptor (bottom right), showing that each TEB surface only interacts with one side of the spacer unit. The bottom left spectra confirm that the ureas are pointed inwards towards the cavity, as no correlation between urea protons $NH_{B/B'}$ and spacer protons are seen.

Studies by NMR were then performed on receptor **3** to probe any aggregation or self-association effects, as changes in NMR spectra due to these phenomena are important to characterise, this is so they are not misinterpreted as binding events during later studies.[96] Variable temperature ^{1}H NMR studies were then undertaken. Several resonances sharpened as the temperature increased (Figure 72). The most significant changes were observed for protons t1/t1', t2/2', s1 and s3. This might suggest that these protons are involved in sterically hindering rotation and the higher temperatures provide enough energy to exceed this barrier to rotation. Such steric effects may be the reason for the rigid conformation where the ureas are directed inwards towards the cavity. A dilution study of receptor **3** was also performed using ^{1}H NMR, and it was found that there was no change in the spectra over the concentration regime 1-0.05 mM. This would imply that the receptor is monomeric and does not aggregate over this concentration range at least.

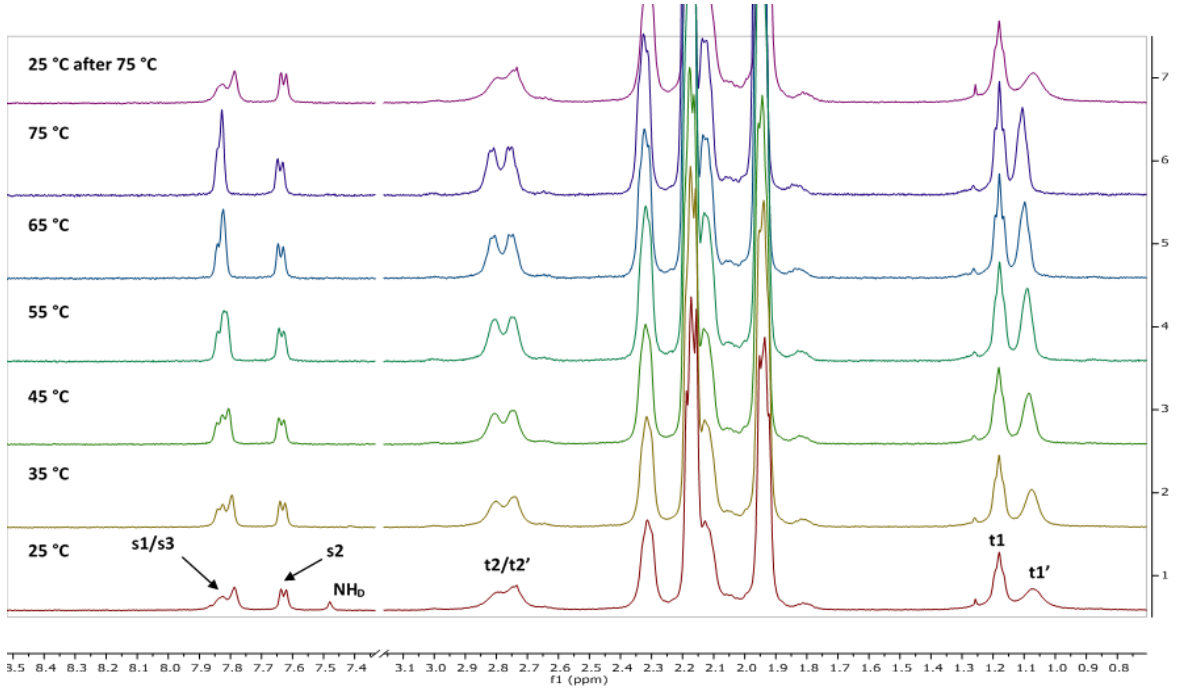


Figure 72 Partial ^{1}H NMR spectra (500 MHz, D_2O, pH 7.4) showing receptor **3** at various temperatures. Receptor signals sharpen upon heating – most notably for t1/t1', t2/t2', s1 and s3. Chemical shifts (δ, ppm) relative to t1 (1.18 ppm). Signal for t3/t3' omitted for clarity due to overlap with residual solvent peak that varied in chemical shift upon heating.

The thermal stability of the receptor was also tested, initially by heating **3** (in D_2O) in a sealed vessel. No changes were observed by NMR spectroscopy until 100 °C, whereupon decomposition was observed (Figure 73). The new peaks were much more resolved than that of the receptor at room temperature and it was found the spectra remained unchanged upon cooling to 25 °C. It became apparent that hydrolysis of the some ureas had occurred to give an acyclic molecule, as signals for the TEB surface were still visible. However, heating **3** to 150 °C in the solid state appeared to have no effect, with no evidence for decomposition. Such stability is desirable for sterilisation purposes, if the receptor were incorporated into a medical device.[97] Receptor **3** was also tested for cytotoxicity towards HeLa cells, studies of which were performed by Hongyu Li. It was found that incubation of the HeLa cells for 18 hours, with the receptor present at a range of concentrations (0.01 mM – 1 mM), gave no indication of any toxic effects. Low or non-toxicity is critical towards medical applications, and these initial studies bode well for the potential implementation of **3** into medical devices that interface with the body. It is speculated that receptor **3** is not cell permeable, due to its polycarboxylate solubilising groups, and thus is not cytotoxic.

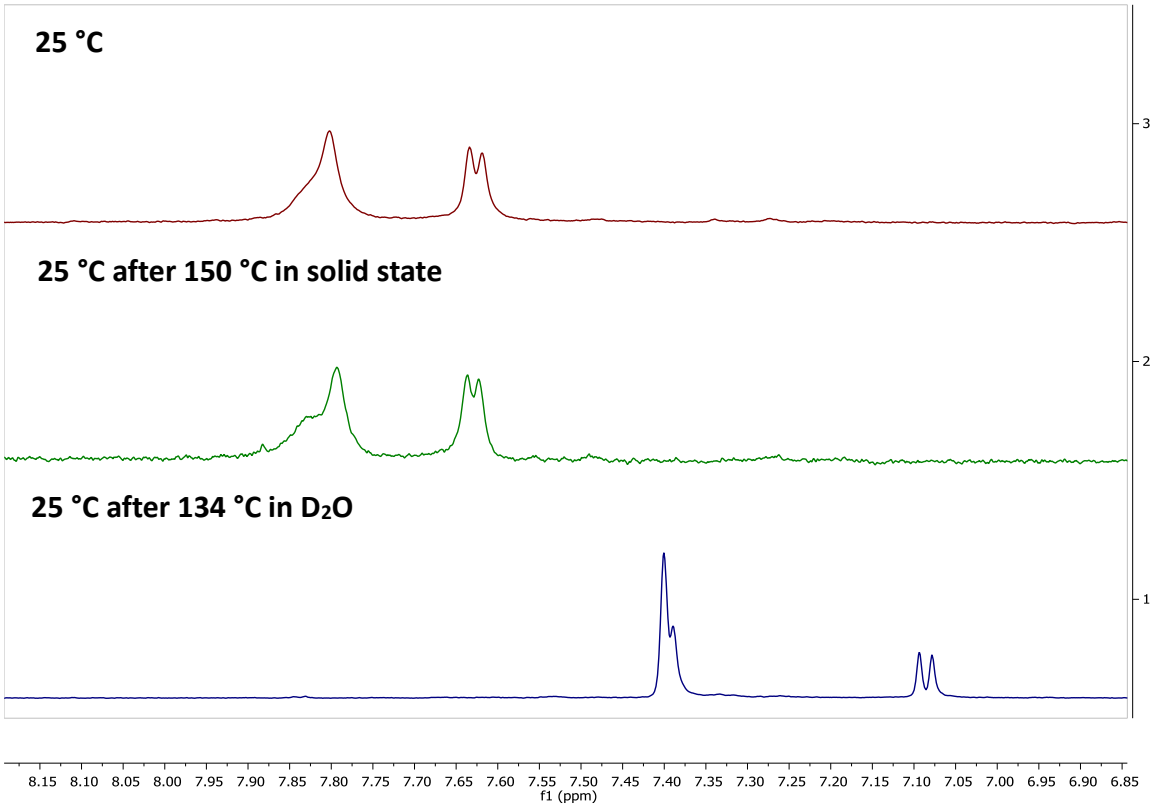


Figure 73 Partial ^{1}H NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) of **3** at 25 °C (top), and after heating at 150 °C for 1 hour in the solid state (middle), showing no signs of decomposition. Heating to 134 °C in D_2O however showed complete decomposition of the receptor (bottom).

5.4 Binding studies

Given that receptor **3** was designed with β-D-glucose **4** as a target substrate, initial binding studies with **3** and D-glucose **4** were performed by ^{1}H NMR in D_2O. Both stock solutions of receptor **3** and D-glucose **4** were confirmed to be at pH 7.4 (with 10 mM phosphate buffer solution) to avoid variations in pH giving misleading changes in the NMR spectrum. The solution of D-glucose was allowed to reach equilibrium of both α and β anomers by dissolution in the titration medium (here D_2O) 24 hours beforehand, with the equilibrated anomeric ratio verified by ^{1}H NMR. A solution of D-glucose (~100 mM), with added receptor **3** (~0.3 mM), was titrated into a solution of **3** (~0.3 mM). This is so that the concentration of receptor **3** remains constant, as guest concentration increases, throughout the titration and makes extraction of the binding constant from the data much simpler. A modest K_a of ~1000 M^{-1} was estimated initially for the complexation of D-glucose with **3**, and the planned titration would have given a good number of data points if this were the case. However, it became rapidly apparent that the binding constant was much higher. The symmetrical NMR spectrum of free receptor **3** became very convoluted (especially in the aromatic region) upon the initial addition of D-glucose,

which gave a final glucose concentration of 2 mM (or roughly 5-6 equivalents). Further additions of D-glucose appeared to have no effect on the spectrum, implying the receptor was most likely fully bound with guest after the first addition of D-glucose. No determination of the binding constant was possible at this point, although rough calculations based on concentrations of host and guest at this period in the titration estimated the K_a to be ~10^4 M^{-1}.

The NMR titration was then repeated where the amount of glucose titrated into receptor was much reduced (Figure 74). The NMR titration spectra indicated that the binding of glucose to **3** was occurring with slow exchange relative to the NMR timescale (see Figure 4, Chapter 1). Integration of part of the aromatic region that relates to host guest complex (Int_{HG}) relative to the whole aromatic region (Int_H + Int_{HG}) can be used to calculate the total concentration of host guest complex ([HG]) at each point in the titration. This can then be used to calculate concentrations of free host ([H]) and guest ([G]) and thus determine the binding affinity. Applying this to each point in the titration of glucose with **3** and fitting the collated data to a 1:1 binding model (one guest per one host) gave a binding constant of K_a ~ 18,000 M^{-1}. This affinity for D-glucose by a synthetic lectin using non-covalent interactions in water is unprecedented, with the previous record being two orders of magnitude lower – receptor **34** with K_a ~ 250 M^{-1} (see Figure 19, Chapter 1).[40]

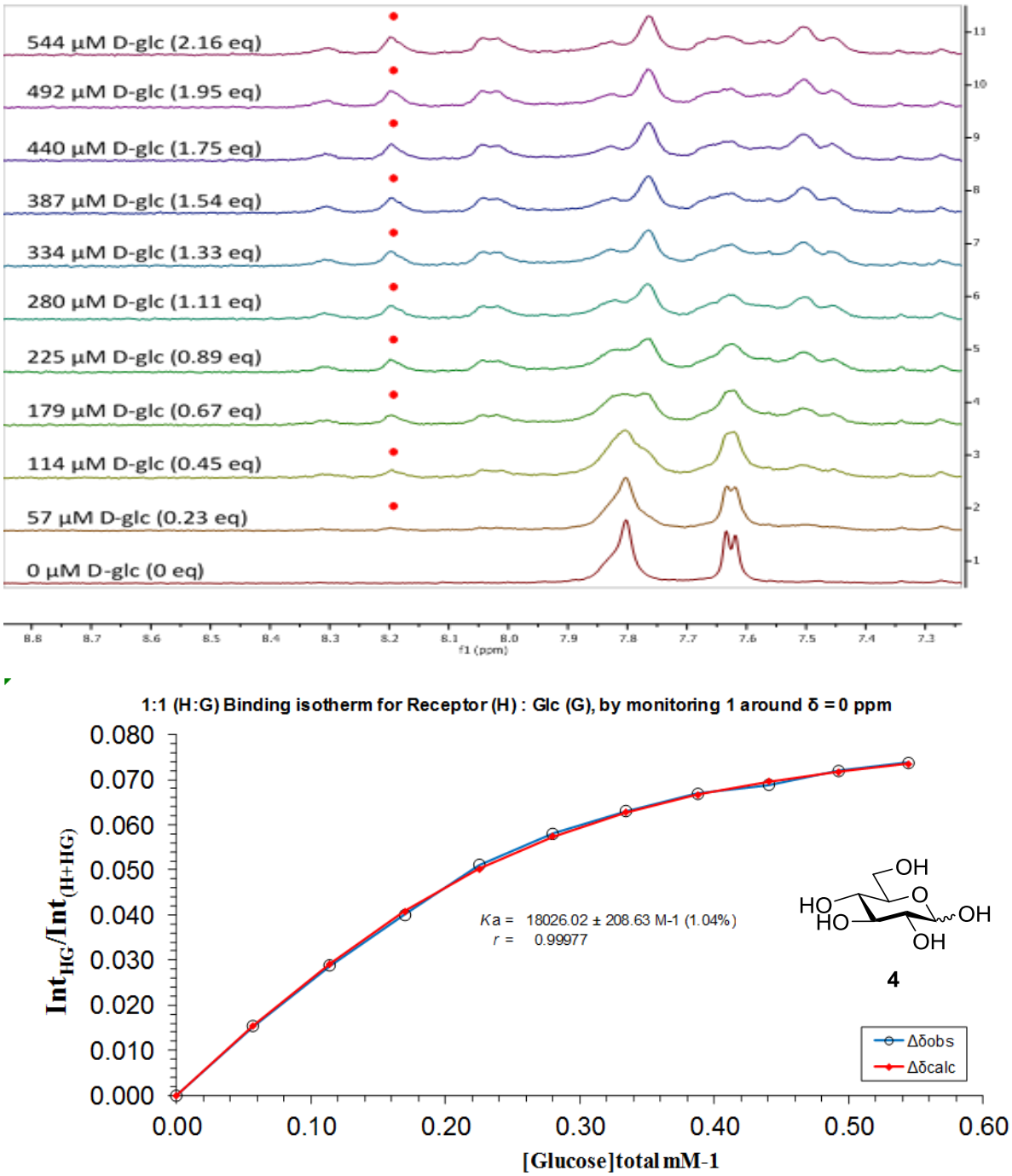


Figure 74 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.25 mM) titrated with a combined solution of D-glucose **4** (9.6 mM) and receptor **3** (0.25 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with slow exchange on NMR timescale. Integrations of the peak at 8.19 ppm (denoted with •) versus the region 8.35-7.39 ppm were plotted against D-glucose concentration (mM). The calculated values for the integrals are overlaid with the observed values, giving K_a = 18,026 ± 208 M^{-1} (1.04%).

The binding affinity for D-glucose to **3** was then corroborated using isothermal titration calorimetry (ITC) experiments (see Chapter 4 for a more in-depth explanation of this technique). Titration of D-glucose (7.5 mM) into receptor **3** (0.13 mM) using ITC gave a very large exotherm upon addition of glucose (Figure 75). Subtraction of the heat of dilution from the titration gave the change in enthalpy for binding, which was plotted against concentration to give a curve. Fitting this curve to a 1:1 binding model yielded a binding affinity for D-glucose and **3** of K_a ~ 18,000 M^{-1} – essentially the same as determined by NMR titrations. The thermodynamic results obtained for D-glucose and **3** shows that the main contribution of binding appears to be enthalpically based, with a smaller contribution from the entropy of binding. This is most likely due to the strong hydrogen bonding interactions formed between glucose and the receptor being the main driving force for complexation. Large increases in entropy upon binding are usually indicative of many water molecules being displaced from the binding site or an increase in receptor/substrate flexibility upon binding. Neither would appear likely in the case of **3**, given its very rigid structure and small cavity size, and this was reflected in the thermodynamic results. The large enthalpic contribution towards binding is too large to be solely a result of CH-π interactions between the receptor **3** and substrate – the enthalpy gain for binding interactions between one carbohydrate and one aromatic is typically ~8 kJ mol^{-1}.[98] Therefore, it can be speculated that a large part of the driving force of the binding affinity is hydrogen bonding between the ureas and the sugar hydroxyls, which is all the more impressive given the competitive aqueous solvent.

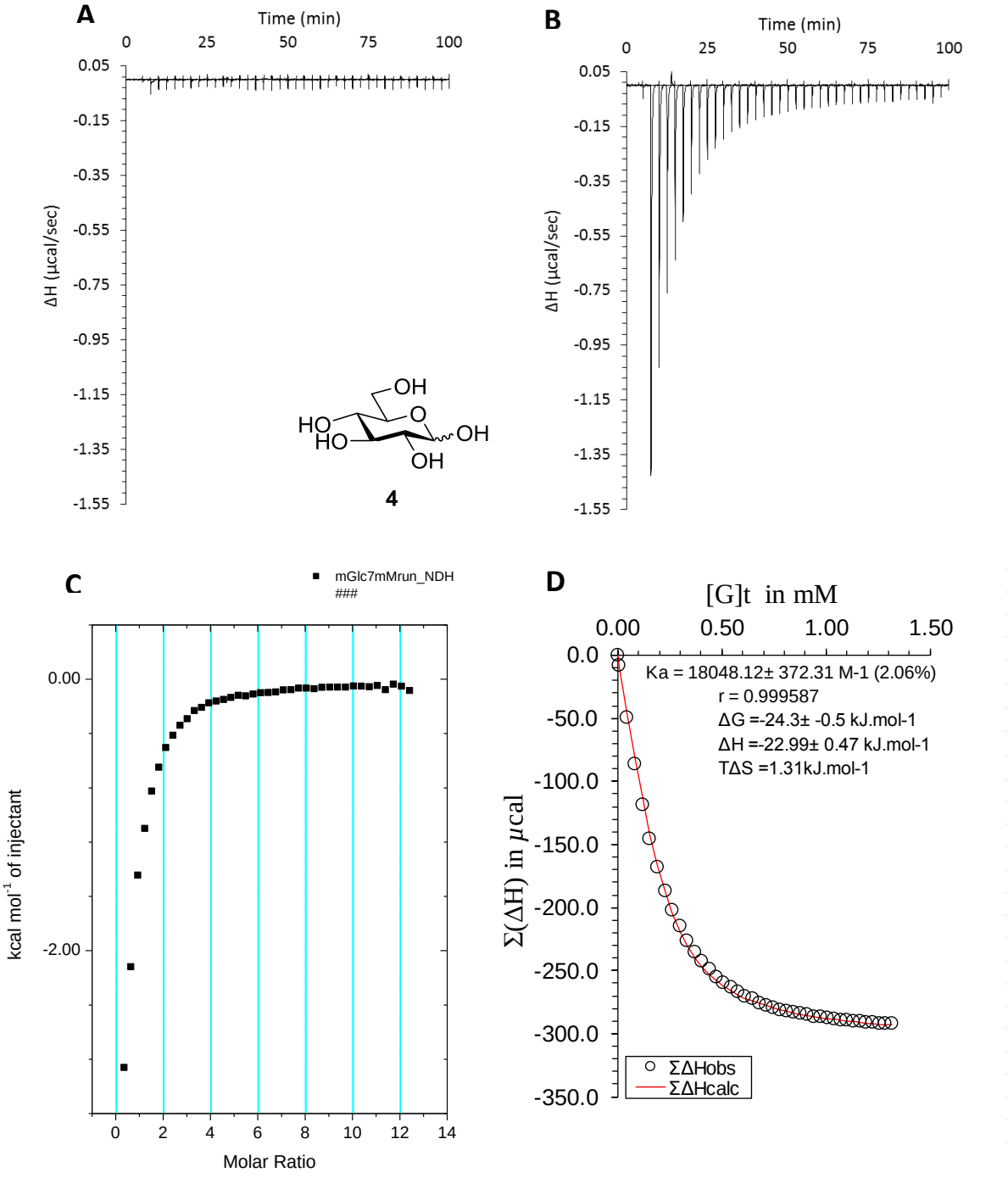


Figure 75 ITC binding results for receptor **3** (0.2 mM) titrated with D-glucose **4** (7.5 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor 2); C) shows the plotted change in enthalpy vs molar ratio and D) shows the fit calculated using a 1:1 binding model (K_a = 18,000 ± 372 M^{-1}).

An NMR titration for receptor **3** and D-glucose **4** was also attempted in a 9:1 H_2O/D_2O solvent medium, to provide information regarding the interactions between the urea NH protons and D-glucose. However, it was found that large broad artefacts were observed during the titration both downfield and upfield from the signal for the bulk water protons (which were suppressed using NMR solvent suppression techniques). These artefacts would appear to change in size and broadness between additions during the titration and would require recalibration of the water suppression between each addition – this proved unpractical given the number of data points required for a NMR titration experiment. The artefacts are speculated to arise from exchange of water in the cavity with the bulk solvent water, as water bound to the receptor would possess a different chemical shift to that of free bulk water. This process can be considered comparable to the effect MRI contrasting agents have upon water, whereby water bound to the contrasting agent is perturbed enough to exhibit a distinct signal compared to the bulk water.[99] Given that this phenomenon would be unavoidable, it was deemed not practical to pursue an NMR titration in this medium.

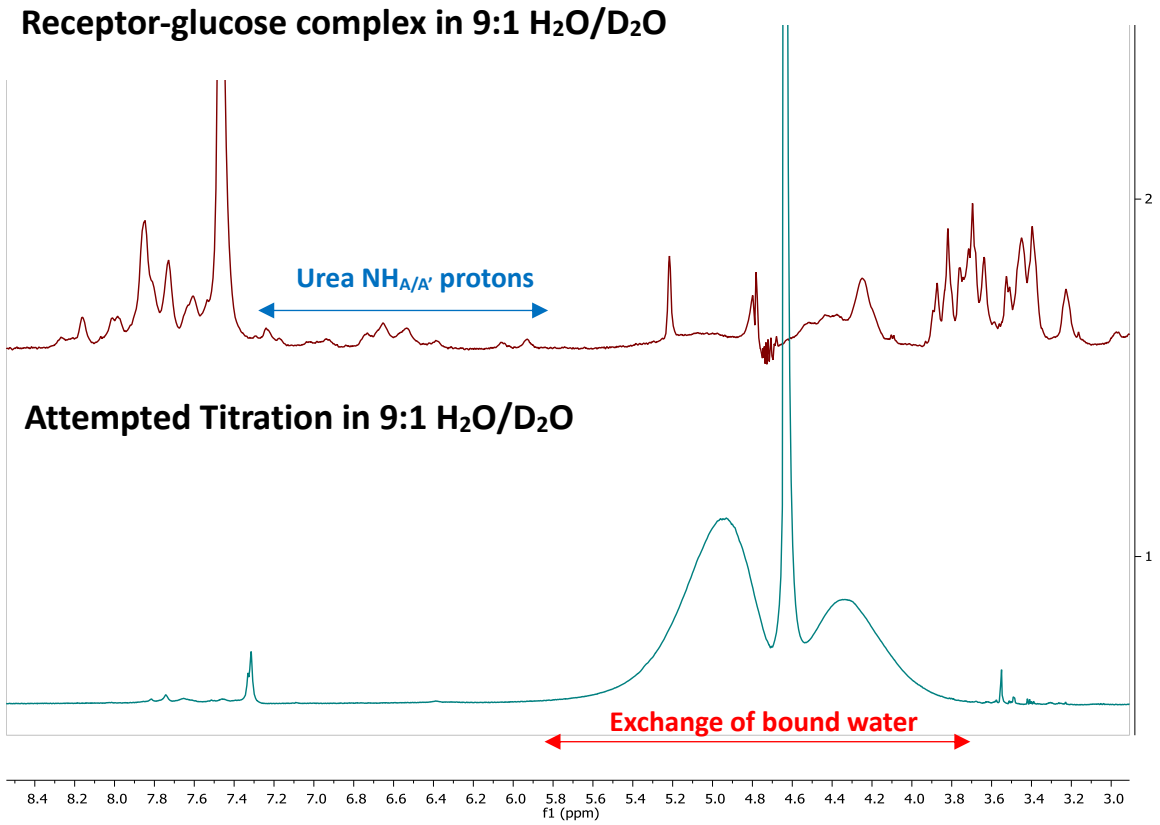


Figure 76 Partial ^1H NMR spectrum (600 MHz, pH 7.4, 298 K) of **3** saturated with D-glucose in 9:1 H_2O/D_2O (top), where the complex was preformed. The bottom spectrum is part of an attempted NMR titration of D-glucose into **3** in 9:1 H_2O/D_2O, where large broad artefacts are observed δ ~3.6-5.8 ppm. These artefacts are speculated to arise from slow exchange of bound water molecules in the receptor binding site with bulk solvent water molecules.

However, preformation of the saturated receptor and glucose complex in 9:1 H_2O/D_2O was predicted to feature very limited (if any) exchange of water in and out of the cavity, and therefore these bound water artefacts would not be present. Indeed, such a result was observed, with no broad artefacts present adjacent to the bulk water signal. The NMR spectrum of the complex was now more convoluted than when obtained in pure D_2O, as all NH protons will be visible. Part of this complexity is presumably due to the fact that all urea NH protons (12H) are now inequivalent when bound to an unsymmetrical guest (i.e. D-glucose). Given the asymmetry of the receptor itself with regards to the 'horizontal axis', there are two main possible binding orientations of glucose with the receptor. This further complicates the number of individual urea NH protons in the NMR spectrum for the complex – potentially generating 24 unique urea NH environments. This can be seen to some effect for the benzylic urea NH protons, which appear to occupy a chemical shift range of > 1 ppm in the ^{1}H NMR spectrum of the complex (Figure 76).

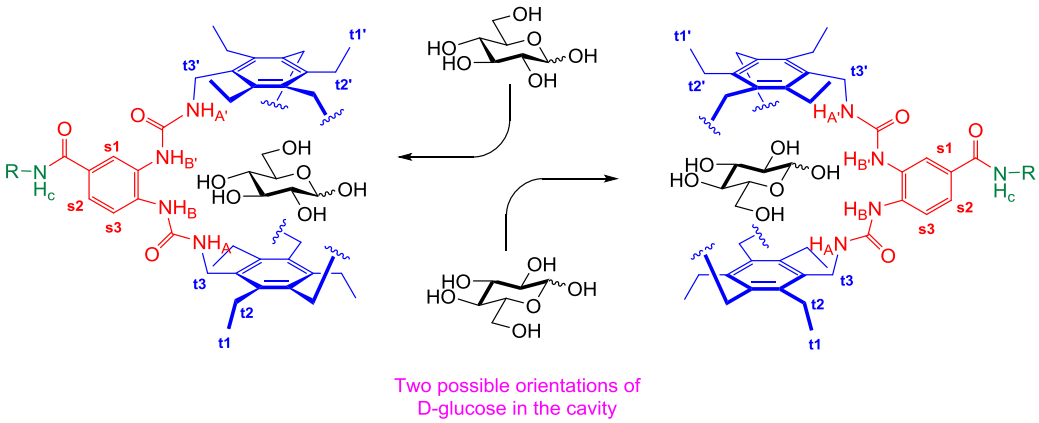
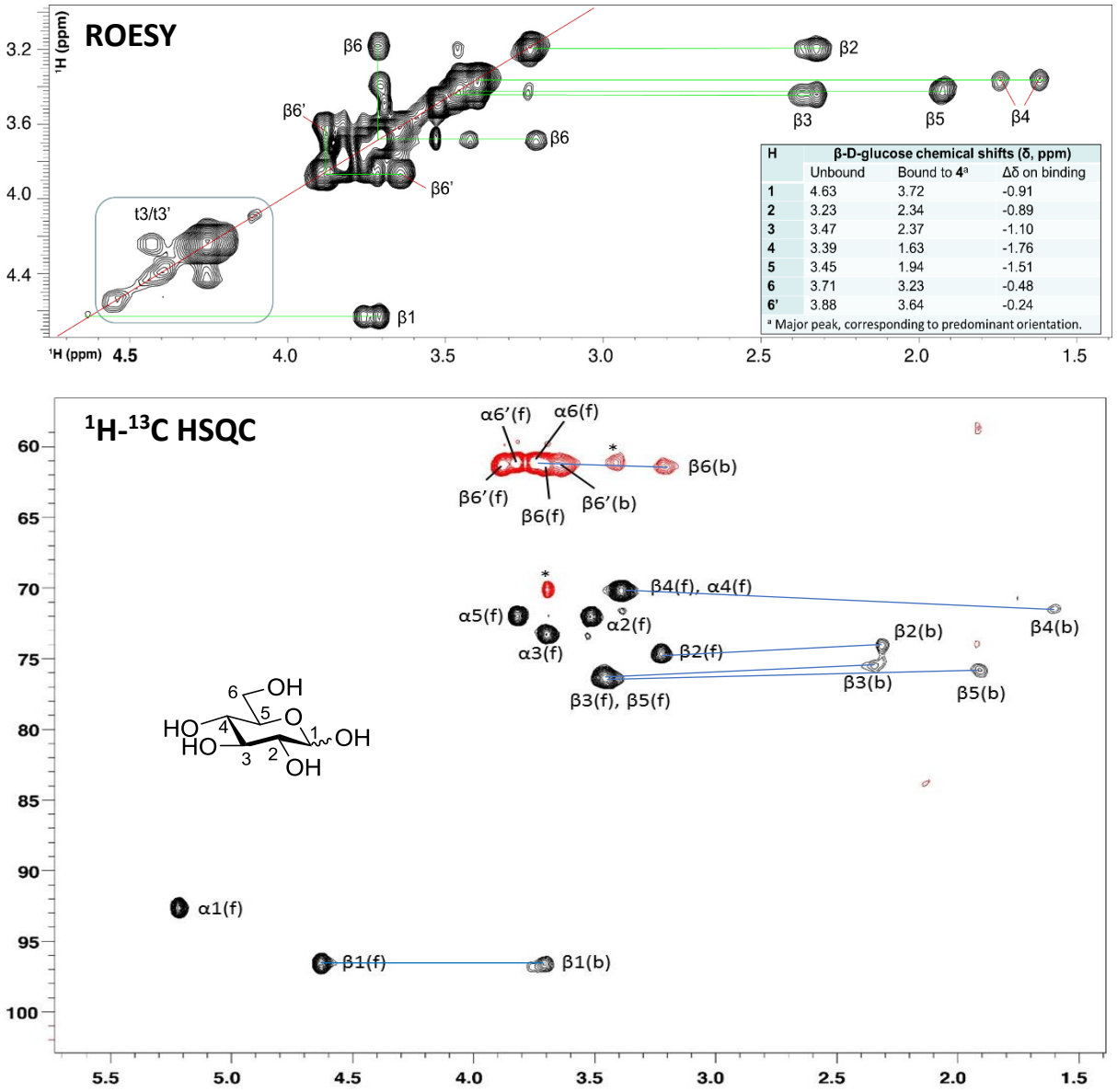


Figure 77 Schematic of two different possible orientations of D-glucose inside the receptor cavity. Each complex would lead to a potential of 12 different urea NH environments, giving 24 different urea NH environments for both complexes combined. Potential evidence for this can be seen in Figure 76.

One of the advantages of a host guest system exhibiting slow exchange kinetics upon binding, as seen for **3** and D-glucose, is that a large amount of structural information of the complex can be garnered using 2D NMR experiments. Analysis of the receptor and glucose complex was then performed using 2D TOCSY, NOESY and ROESY ^{1}H NMR experiments. These spectra further exemplified the asymmetry of the receptor-glucose complex, with many different protons environments visible. Signals relating to bound glucose were obscured by receptor protons (most notably the dendrimer solubilising groups) in the 1D ^{1}H NMR, but could be observed using these 2D NMR methods. A ROESY spectrum in D_2O very clearly indicated chemical exchange peaks between free and bound glucose, with signals for the bound form shifted upfield approximately 1.5 ppm (Figure 78). These signals are shifted upfield due to the receptor shielding the bound substrate from the magnetic field, and are most notable for positions 3-CH, 4-CH, and 5-CH of β-D-glucose – suggesting these regions of the substrate are located most deeply

within the cavity. Upfield movements of the 6-CH_2 were relatively small however, which would imply that the CH_2OH extends outwards from the cavity – an observation that is consistent with molecular modelling of **3** and glucose. Also prevalent were two species of bound glucose, with two distinct signals observed for the bound substrate (most notably for 4-CH). This was consistent with the prediction of two major orientations of glucose within the receptor cavity (Figure 77). No signals corresponding to the α-anomer of glucose were observed under these conditions (mixing time of 100 ms). However, use of longer mixing times (>500 ms) did reveal signals for the bound α-anomer. No quantitative deductions of selectivity between the α and β anomer could be concluded from these data, as the exchange peaks observed are dependent on exchange rates of glucose in and out of the receptor cavity and not from actual populations of bound and unbound species. Determination of selectivity between the anomers was determined using ^{1}H-^{13}C HSQC however, with signals for only β-D-glucose bound to **3** visible. The intensity of signals for bound β-D-glucose relative to baseline noise was measured to be 6:1 (signal:noise). If the signals for bound α-D-glucose are at most the same intensity of the baseline noise, then it can be presumed that the selectivity of **3** for β and α-D-glucose is at least 6:1 (β:α). The selectivity could potentially be higher however, but would require a much more concentrated sample or longer acquisition time – neither of which were feasible at this time.



H	β-D-glucose chemical shifts (δ, ppm)		
	Unbound	Bound to **4**[a]	Δδ on binding
1	4.63	3.72	-0.91
2	3.23	2.34	-0.89
3	3.47	2.37	-1.10
4	3.39	1.63	-1.76
5	3.45	1.94	-1.51
6	3.71	3.23	-0.48
6'	3.88	3.64	-0.24

[a] Major peak, corresponding to predominant orientation.

Figure 78 Partial ROESY (^{1}H-^{1}H correlation, through space, top spectrum) NMR spectrum (600 MHz, D_2O, pH 7.4, 298 K) of receptor **3** (2 mM) with D-glucose **4** (5 mM, 2.5 equivalents). Chemical exchange peaks are observed between the D-glucose bound to the receptor (whose signals appear more upfield) and free D-glucose. Only signals relating to bound β-D-glucose are labelled, as exchange peaks are observed for only the β anomer under these conditions (mixing time of 100 ms). Exchange peaks for the α anomer were observed at much longer mixing times (>500 ms). Partial HSQC (1H-13C correlation, one bond, bottom spectrum) NMR spectrum of the same sample. The signals for "free" and "bound" states of D-glucose are shown in black for CH protons and red for CH2 protons, and are labelled as for the ROESY spectrum. NMR signals corresponding to bound D-glucose are observed for the β anomer only, indicating an expected preference for binding the β anomer. The ratio of intensity of signals for bound β-D-glucose and baseline noise is measured as 6:1 (signal:noise). Assuming the bound signals for the α anomer are at most the same intensity as the baseline noise, then it can be assumed the selectivity for β and α-D-glucose is also at least 6:1 (β:α) but could potentially be higher.

The receptor-glucose complex was also investigated using 2D NOESY NMR, using 9:1 H_2O/D_2O as solvent to enable visualisation of the NH protons (Figure 79). This was to allow observation of interactions between the urea NH protons and bound glucose. Indeed, such interactions were observed but the sheer complexity of the spectra obtained meant that characterisation of individual interactions was non-trivial. Interactions could be grouped by similarity however, and each of the urea ($NH_{A/A'}$ and $NH_{B/B'}$) protons showed distinct interactions with D-glucose. The inequivalence of the urea protons in the complex was now very apparent, with $NH_{A/A'}$ in particular displaying signals over a chemical shift range of δ >1.5 ppm. Also evident were interactions between the ethyl groups (protons t1/1' and t2/2') and glucose. This may suggest some hydrophobic interactions are occurring between the ethyl groups on the receptor surfaces and the axial CH protons of glucose – this may effectively increase the size of the hydrophobic surface of the receptor and thus increase binding affinity.

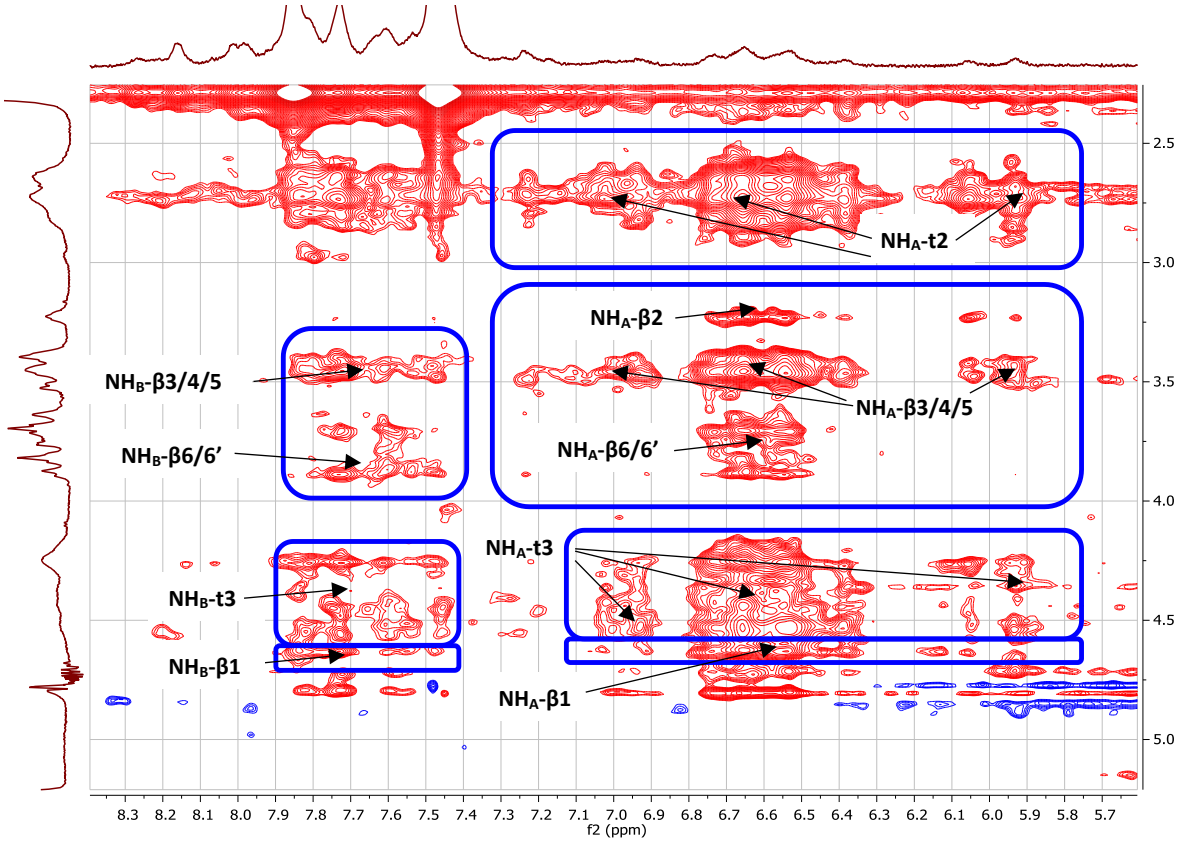


Figure 79 Partial NOESY (^1H-^1H correlation, through space, 500 ms mixing time) NMR spectrum (600 MHz, pH 7.4, 298 K) of **3** and D-glucose in 9:1 H_2O/D_2O. Correlations between urea NH and glucose protons highlighted with blue boxes and labelled accordingly.

Given that receptor **3** displays a clear preference for binding β-D-glucose over α-D-glucose, it was deemed imperative to determine whether **3** influences the equilibration between anomers. The previous association constants (K_a) determined for **3** and glucose are 'apparent association constants', as D-glucose is effectively composed of two different substrates (i.e. α and β anomers) with largely differing affinities to **3** based on previous binding data (Figure 78). These obtained binding constants are calculated under the assumption that the anomeric ratio is unperturbed by binding of β-D-glucose to the receptor. However, the anomeric ratio could be affected by this through the sequestering of β-D-glucose in the receptor causing the amount of free β-D-glucose in solution to be lowered. As a consequence of this, some α-D-glucose then is converted to β-D-glucose to compensate for this 'removal' of β-D-glucose from the equilibrium. This newly generated β-D-glucose is then bound by receptor **3**, and this process repeated until almost all α-D-glucose is converted to the β-anomer and thus bound by **3**. This scenario would effectively result in the amount of binding guest (i.e. β-D-glucose) to steadily increase between additions during the titration, making the concentration of binding guest unknown at each point – this would produce inaccurate calculated binding constants, as the value of binding affinity is directly proportional to guest concentration. This scenario would be true if equilibration between free anomers in solution is fast relative to binding to the receptor.

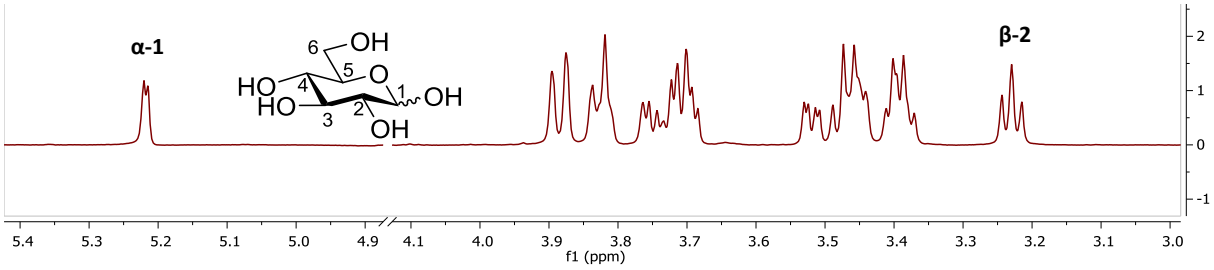


Figure 80 Partial ^{1}H NMR spectrum (600 MHz, D_2O, pH 7.4, 298 K) of D-glucose. The peaks corresponding to the protons at position CH-1 for α-D-glucose and CH-2 for β-D-glucose are labelled accordingly. These protons were selected to determine (using integration) the α:β anomeric ratio during NMR studies into whether receptor **3** affects the anomeric equilibrium.

To determine whether receptor **3** does indeed influence the anomeric equilibrium, a series of ^{1}H NMR experiments were undertaken. The titration spectra for D-glucose into receptor **3** were reviewed, with protons at position CH-1 for α-D-glucose and CH-2 for β-D-glucose used to quantify the ratio of anomers using integration (Figure 80). The ratio of unbound anomers was determined at each point in the titration using this method, whereupon it was found that the anomeric ratio remained constant at 64:36 (β:α) past halfway through the titration– the same as for free glucose in water. Integrations of the anomeric ratio earlier on in the titration (i.e. <1 equivalent of D-glucose) were deemed inaccurate due to low intensity and broadening of sugar signals, presumably from exchange with the receptor.

This finding would appear to corroborate with previous literature outlining that if a substrate is composed of two isomers/anomers (i.e. D-glucose), then the ratio of these anomers complexed to a receptor will be constant provided the ratio of free anomers in solution is also constant.[34] This will be true if both anomers in are in sufficient excess relative to the receptor. If this is the case, then the titration will proceed as if only one guest were present, yielding a 'composite' or apparent binding constant. This hypothesis would be consistent with the results observed for receptor **3** binding to D-glucose, which is in approximately 5 fold excess towards saturation of **3** – with NMR studies showing a constant anomeric ratio when possible to determine. Attempts to plot the titration data assuming only β-D-glucose binds (so effectively only 64% of the total concentration of D-glucose used) results in data that does not fit to a binding model, with no association constant determinable. This would suggest that the apparent binding constant must be resulting from binding of both anomers to receptor **3**, albeit with a preference for the β-anomer.

To further investigate this, and see if receptor **3** could accelerate equilibration between anomers of D-glucose (effectively acting as an enzyme – an 'anomerase'), another NMR experiment was undertaken. Pure α-D-glucose was dissolved in D_2O and 1D ^1H NMR spectra recorded over time until equilibrium was reached, with the ratio of anomers deduced using integration as described earlier (Figure 80). This was then repeated in the presence of receptor **3**, with the ratio of anomers again calculated at certain intervals over time. It was found, however, that the receptor appeared to have no effect on the rate of equilibration of α-D-glucose to β-glucose. Plotting the ratio of α and β anomers against time gave very similar linear curves, with rate of equilibration appearing constant with or without receptor present. In light of these results, it was therefore concluded that receptor **3** does not influence the equilibration between α and β-anomers of D-glucose, and that the binding constants obtained earlier are indeed accurate.

While high affinities for substrates can be difficult to obtain (especially for carbohydrates in water), the ultimate challenge is selectivity for a target substrate over all others – especially substrates with similar structural properties. This is particularly difficult for a molecule trying to distinguish between carbohydrates, as inversion of the stereochemistry of a single hydroxyl can result in an entirely different monosaccharide – often with unique and important biological properties. A receptor that is unable to distinguish between two carbohydrates would not be useful in a medical context as such off-target binding would give false results if incorporated into a biological sensor; or even outright dangerous if the binding disrupted biological pathways of key carbohydrate substrates. Given the excellent affinity of hexaurea **3** for D-glucose **4**, quantification of its selectivity for D-glucose **4** over other substrates was critical if **3** was to ever function in a medicinal context. The selectivity of receptor

3 was then surveyed by performing binding studies with other substrates, including similar carbohydrates, by ^{1}H NMR and ITC titrations as for D-glucose **4**.

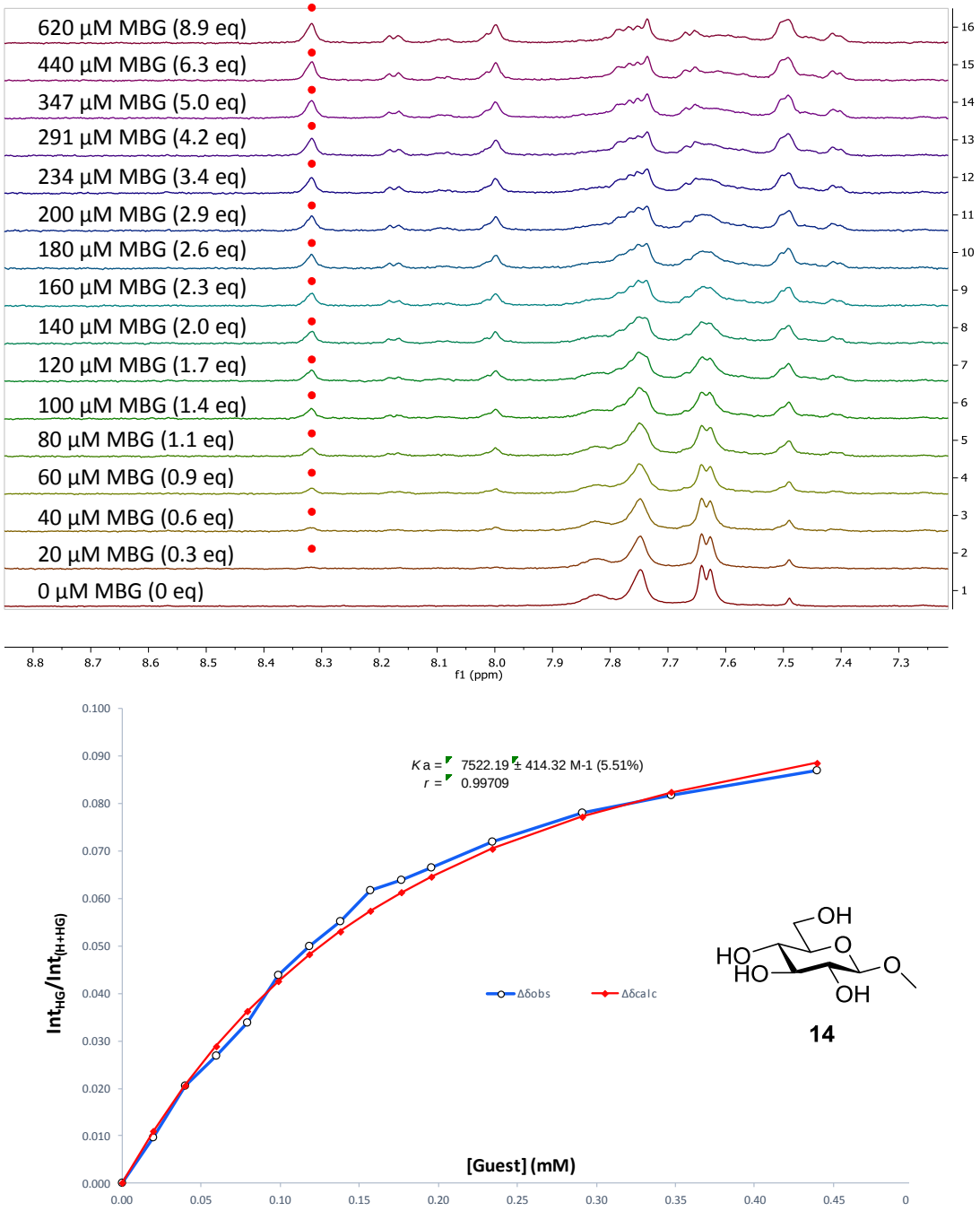


Figure 81 ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.07 mM) titrated with a combined solution of methyl β-D-glucoside **14** (10 mM) and receptor **3** (0.07 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with slow exchange on NMR timescale. Integrations of peak at 8.31 ppm (denoted with •) versus region 8.36-7.36 ppm were plotted against guest concentration (mM). The calculated values for the integrals are overlaid with the observed values, giving Ka = 7522 ± 414 M^{-1} (5.51%)

The first substrate to be tested was methyl β-D-glucoside **14**, as this is an all equatorial carbohydrate (like D-glucose) but only exists as the β-anomer due to methylation of the anomeric OH – this prevents ring opening and thus possible formation of two isomers upon re-cyclisation (as for reducing sugars such as D-glucose). In past receptors from the Davis group, **14** has exhibited increased affinities over D-glucose due to the increased hydrophobic nature of the methyl group vs the free hydroxyl in D-glucose. For **3**, however, the opposite trend was observed – whereby the affinity of **14** and **3** is lower than that of **3** and D-glucose **4**. NMR titrations for **14** and receptor **3** suggested slow exchange binding kinetics (as for D-glucose), although signals relating to the sugar-receptor complex were more defined for **3** than for D-glucose **4** – potentially due to there being only one anomer of **14** unlike D-glucose **4**. Integration of the spectra and fitting the data to a 1:1 binding model gave an association constant of K_a ~7500 M^{-1}, almost 2.5 times weaker than for D-glucose **4** (Figure 81). This binding affinity was also confirmed by ITC. The observed reduction in affinity is presumed to be due to the increased steric effects of methyl β-glucoside **14** compared to D-glucose, with the anomeric methyl group making the substrate larger and a poorer fit for the cavity.

A similar result was also obtained for D-glucuronic acid **168**, which displayed a binding affinity of K_a ~5400 M^{-1} using ITC. No binding constant for **168** was able to be determined by NMR titration as the binding of receptor and substrate appeared to operate under 'intermediate exchange' kinetics (Figure 82). This resulted in severe broadening of receptor signals as **168** was added, meaning there was no way to extract any tangible data (through integration or change in chemical shift) to be fitted to a binding model. As for methyl β-D-glucoside **115**, the increased sterics of D-glucuronic acid **168** are small when compared to D-glucose but appear to have a significant negative effect on the binding affinity. This would suggest that receptor **3** is very sensitive to the size of guest substrates – most likely due to its relatively small cavity size.

NMR titrations of receptor 3 with D-xylose **149** and 2-deoxy-D-glucose **169** also implied binding with intermediate rates of exchange – as for D-glucuronic acid **168**. Again, this meant that no binding constant was determinable using the NMR data. Binding affinities were measurable using ITC however, with D-xylose **149** having an association constant of K_a ~5800 M^{-1}. Most surprising however was that 2-deoxy-D-glucose **169** displayed a binding constant of K_a ~725 M^{-1}, which is approximately 25 times weaker than D-glucose (Figure 83). These reduced affinities are most interesting as both substrates feature reduced sterics over D-glucose **4** (with removal of certain hydroxyls compared to D-glucose **4**) and yet remarkably depressed binding affinities. It could therefore be inferred from these results that the hydrogen bonding interactions between the receptor and sugar hydroxyls are thus critical to the strength of binding affinities. The 2-OH hydroxyl is clearly critical to the binding affinity, given the

dramatic reduction in binding affinity for 2-deoxy-D-glucose **169** when compared to D-xylose **149** and D-glucose **4**.

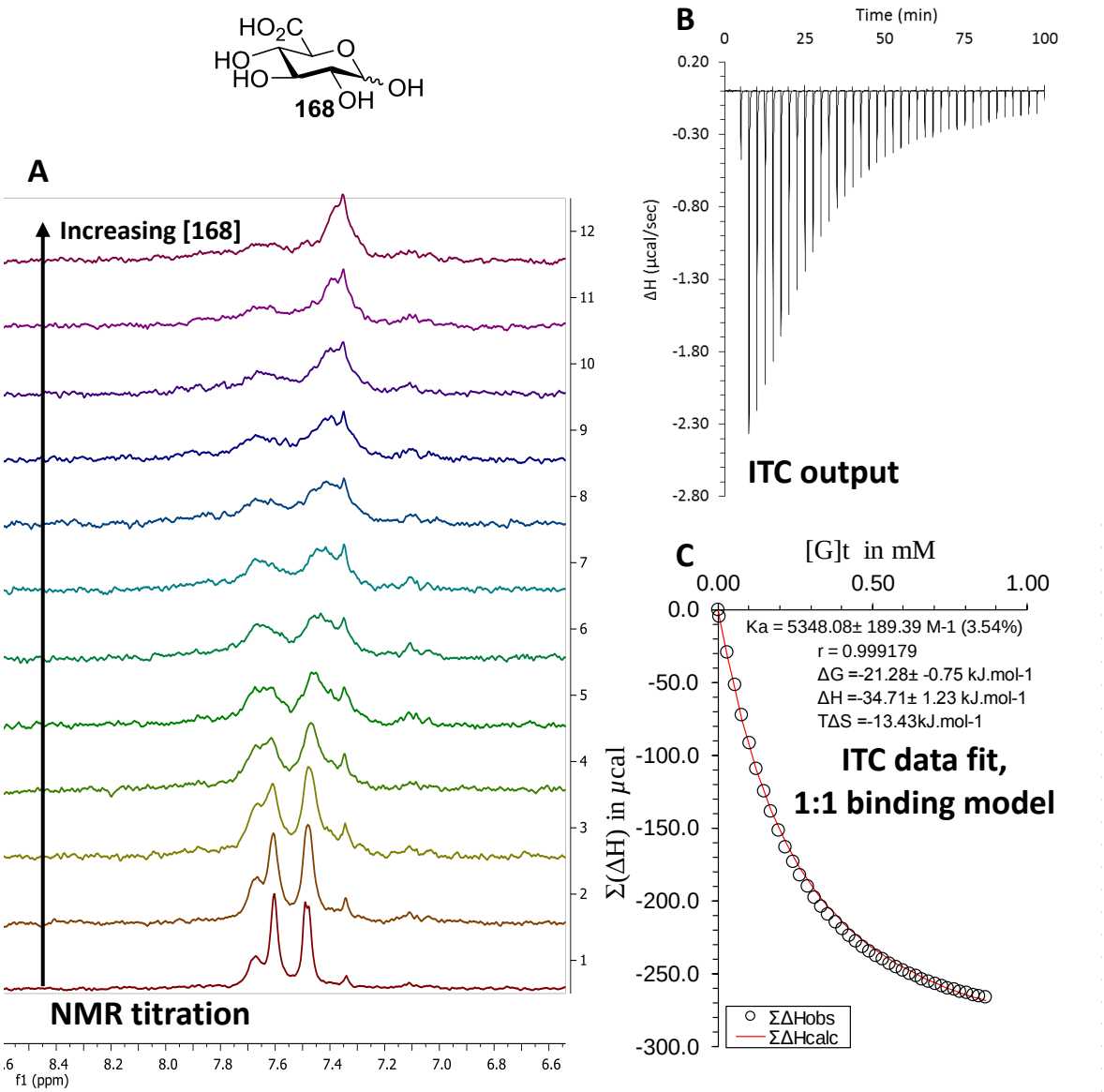
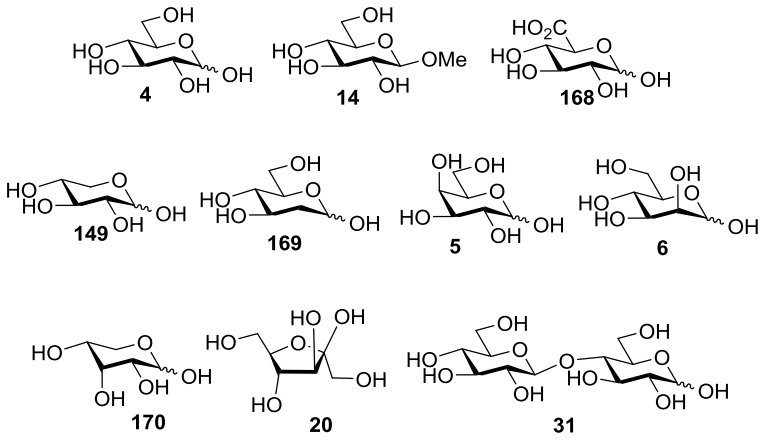


Figure 82 Binding results for receptor **3** and D-glucuronic acid **168**. Attempted NMR titration (600 MHz, D_2O, pH 7.4, 298 K) (A) showed extreme broadening of receptor peaks upon addition of **168**, which imply binding with intermediate rate of exchange – no binding constant was thus determined from these data. Binding constant was determined using ITC (B), which gave a K_a of ~5400 M^{-1} when the data was fit to a 1:1 binding model (C).

Inversion of a single hydroxyl (relative to D-glucose) also appeared to significantly weaken binding. This was demonstrated with NMR and ITC titrations for D-galactose **5** and D-mannose **6**, both of which were determined to have binding affinities to **3** of K_a <200 M^{-1} – two orders of magnitude lower than for D-glucose **4** (Figure 83). This is equal to receptor **3** having a D-glucose vs D-galactose/D-mannose selectivity of >100:1, far exceeding previous synthetic receptors which managed a selectivity of up to 10:1. Differentiation between these monosaccharides is considered very difficult, with even natural proteins struggling to differentiate between the different sugars – the lectin concanavalin A binds both D-glucose and D-mannose for example.[11]



Substrate	K_a (M^{-1})	
	NMR	ITC
D-Glucose **8**	18,000	19,000
Methyl β-D-Glucoside **9**	7,500	7,800
D-Glucuronic Acid **10**	n.d.	5,300
D-Xylose **11**	n.d.	5,800
2-Deoxy-D-Glucose **12**	n.d.	730
D-Galactose **13**	130	180
D-Mannose **14**	140	140
D-Ribose **15**	270	220
D-Fructose **16**	50	60
D-Cellobiose **17**	30	30

Figure 83 Binding affinities for various carbohydrates and receptor **3** in water (at pH 7.4, 298 K), showing the selectivity for D-glucose over other similar carbohydrates. n.d. – not determined, due to intermediate rate of exchange upon binding which caused severe broadening of NMR signals upon addition of substrate.

The reduced affinities for D-galactose **5** and D-mannose **6** are presumably due to the axial hydroxyls making both guests less ideal for the receptor cavity for two potential reasons: increased steric interactions make both guests poor fits for the shape of the relatively small cavity, but also that inversion of an equatorial hydroxyl to the axial position severely disrupts the hydrogen bonding interactions from the ureas (see reduced affinity of 2-deoxy-D-glucose **169**). These reductions in binding affinity due to axial hydroxyls are also apparent for D-fructose **20**, which can feature up to 3 axial hydroxyls in the α-pyranose form. This is reflected by the further reduced association constant of Ka ~50 M^{-1} (Figure 83). The fact that D-fructose **20** in solution is also composed of ~30% as the furanose form may also have an effect, as this structural isomer may bind even weaker or not at all. Differentiation between D-glucose **4** and D-fructose **20** has been historically difficult for carbohydrate receptors derived from boronic acids, as arrangements of diols in both substrates are complementary for boronate formation.[31] The selectivity of hexaurea **3** and D-glucose **4** over D-fructose **20** however appears extraordinary, with D-glucose binding more than 350 times more strongly. Substrates **5, 6, 20** and **170** all bound to **3** with a fast rate of exchange relative to the NMR timescale. It could be speculated that these fast binding kinetics may be due weaker intermolecular interactions and thus lower affinities observed for these substrates, but the increased steric factors may also disfavour full encapsulation inside the cavity and thus enable fast exchange between free and bound guests.

A slightly increased affinity (relative to **5, 6** and **20**) was observed for D-ribose **170** however, which also features an axial hydroxyl (3-position). An association constant of ~240 M^{-1} for D-ribose **170** with receptor **3** was determined by NMR and ITC titrations (Figure 83). This increased affinity compared to other sugars with axial hydroxyls may be due to a lack of substitution at the 6-position – alleviating some of the increased sterics of the axial hydroxyl. The affinity of **170** relative to D-glucose **4** however is still very low meaning selectivity of **3** for D-glucose **4** is still excellent. The lowest association constant to be measured for a carbohydrate was for D-cellobiose **31**, which had a measured affinity of K_a ~ 30 M^{-1}, and this was rationalised to be due to the substantially larger size of the substrate (Figure 83). Interestingly, D-cellobiose **31** appears to bind to **3** with slow exchange kinetics on the NMR timescale despite its low affinity. This might imply that while the large size makes the substrate unfavoured for binding to the cavity, the size of the substrate also makes the kinetics of decomplexation slow due to steric interactions.

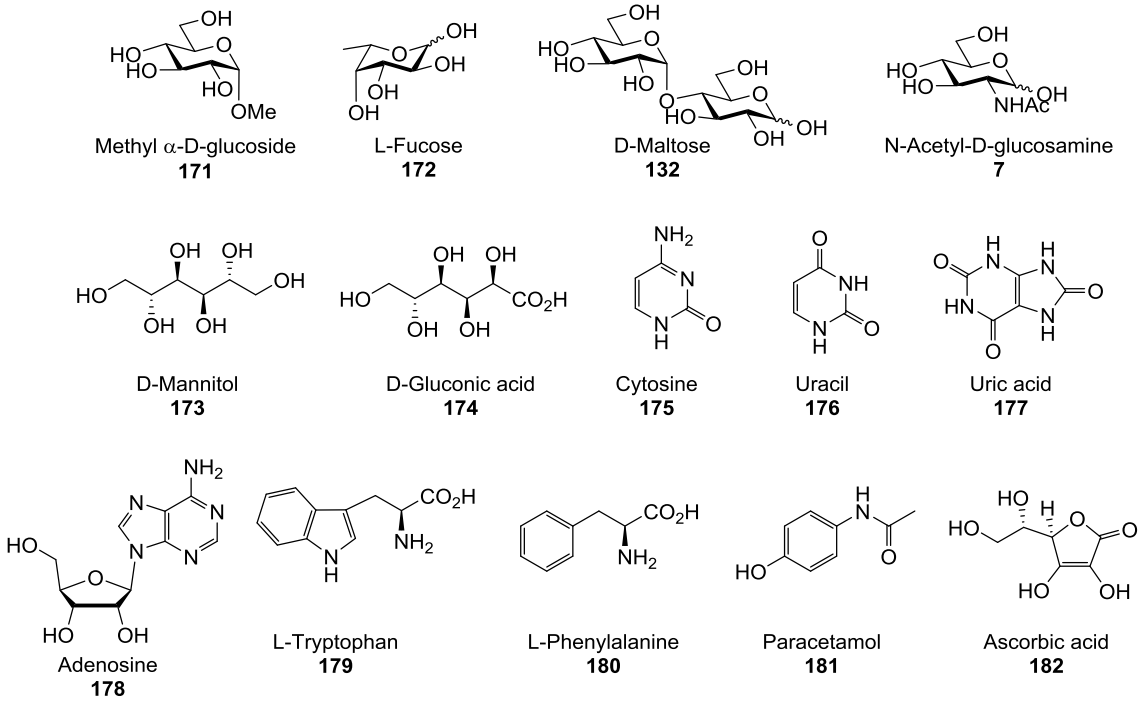


Figure 84 Compounds where binding to **3** was undetectable by ITC titrations, with some also investigated by NMR - although no evidence of binding was observed by this methods as well.

However, many substrates tested for binding affinity by ITC, showed no affinity towards receptor **3** (Figure 84). Given that the affinity of D-cellobiose **31** (K_a ~ 30 M^{-1}) was detectable using ITC, and previous examples in the literature have measured very low association constants using ITC, it is plausible that such low association constants could have been acquired using this technique.[100] Furthermore, distinct differences between the heat of dilution and binding titrations are clearly visible for D-cellobiose **31** but not for these non-binding substrates – where both blank and titration runs are essentially identical (Figure 85). On this basis, there is confidence that these non-binding substrates can indeed be classed as non-binding or have undetectable association constants (i.e. K_a <10 M^{-1}), and therefore selectivity of receptor **3** for D-glucose **4** over these substrates is equivalent to >1800:1.

These non-binding substrates include carbohydrates such as methyl α-glucoside **171**, although this is less surprising given the selectivity observed for β-D-glucose over α-D-glucose for receptor **3**. This coupled with the apparent sensitivity of **3** to sterics (see methyl β-glucoside **14**, Figure 81) can rationalise the lack of affinity seen here for **171**. Carbohydrates L-fucose **172** and D-maltose **132** also exhibit no evidence of binding to receptor **3**, and this can also be rationally explained through unfavourable steric effects as well. Most surprising, however, was the apparent lack of affinity *N*-acetylglucosamine **7** has towards receptor **3**, as it is a carbohydrate with an all equatorial substitution of hydroxyls with one acetamide at the 2-position. Analysing the structure of **7** reveals that the acetamide (position 2) and the CH_2OH (position 6) are 'trans' to each other, and reside at opposite

positions of the pyranose ring. This symmetry is not compatible with the C3-symmetry of receptor **3**, as one of these groups (CH_2OH or acetamide) would always be directed towards one of the urea spacer units and cause unfavourable steric interactions. This incompatible symmetry may mean that **7** simply cannot fit inside the cavity, and would result in very low or no binding affinity. NMR titrations were attempted for these carbohydrate substrates, but no evidence of binding to **3** was observed. Broadening and slight movement of some signals was observed at very high concentrations of guest (>500 equivalents of guest to host), but gave linear straight lines when fit to a 1:1 binding model which implied that very low (i.e. K_a <1 M^{-1}) or no binding was taking place.

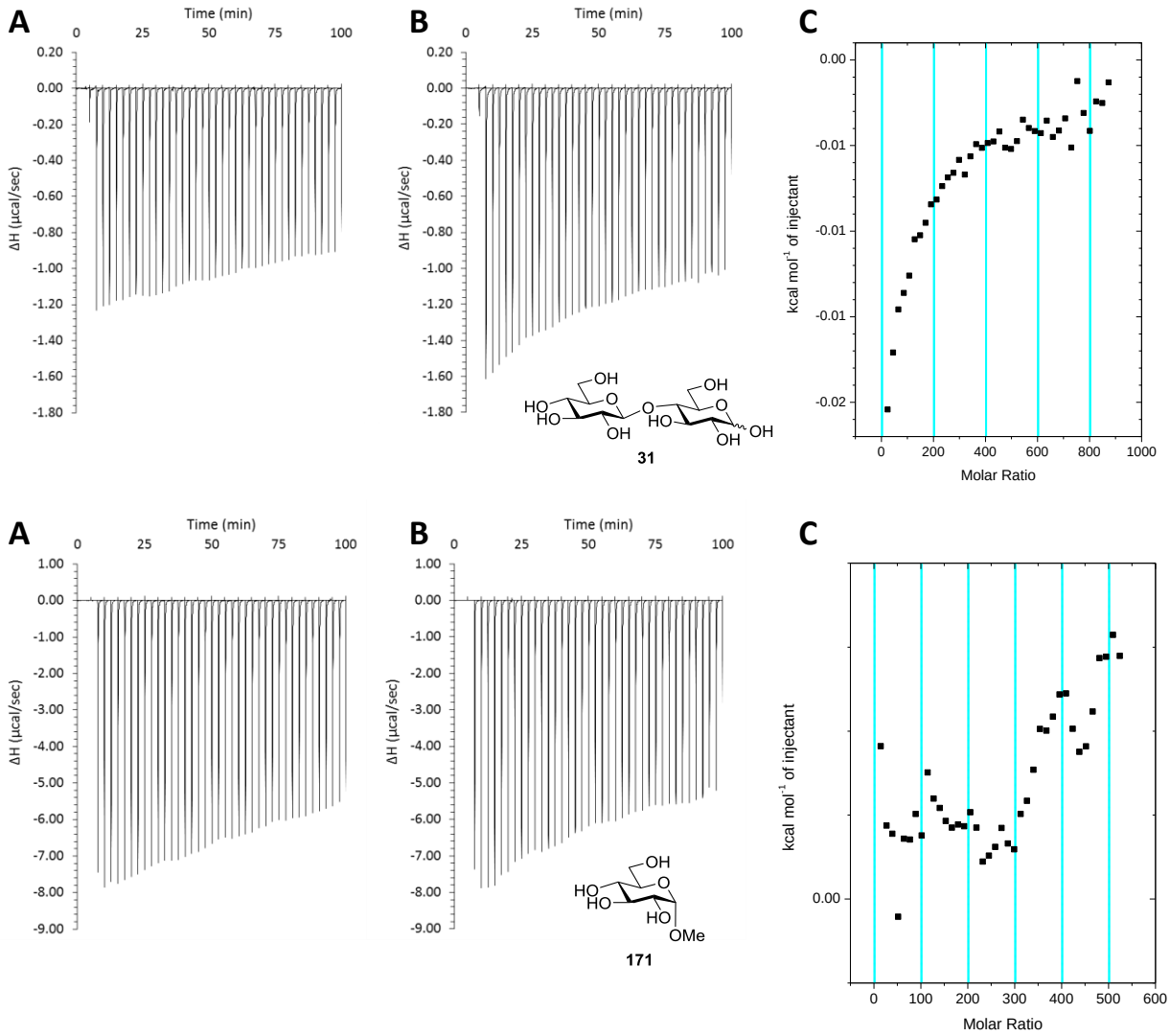


Figure 85 ITC traces for D-cellobiose **31** (top), which show heat of dilution (A), titration run (B) and plotted enthalpy change (ΔH) vs guest concentration (C). The same is displayed for the ITC experiment with methyl α-glucoside **171** (bottom). Distinct difference in ΔH (y-axis) for titration of cellobiose (B) compared to heat of dilution (A). This was not seen for **171**, and thus no typical binding curve seen when plotted (C). These data suggest no binding of **171** to **3**, and is typical for all non-binding substrates.

Linear polyols D-Mannitol **173** and D-Gluconic acid **174** (D-gluconate at pH 7.4) also displayed no evidence of binding to **3**, which infers that the cyclic pyranose structure of D-glucose is essential for binding. No binding was also confirmed by NMR titration. Previous carbohydrate receptors derived from boronic acids tend to complex polyols in general, and have historically had difficulties distinguishing between them and their target carbohydrates. No such issues were evident for receptor **3** however.

Substrates bearing aromatic core structures or pendant aromatic groups, all of which tested are key substrates in human biology, also displayed no evidence of binding to receptor **3**. Nucleobases cytosine **175** and uracil **176** showed no affinity for **3** by ITC, and the same negative result was obtained for uric acid **177** – a major constituent of human blood. Nucleobases guanine and adenine were unable to be tested due to complete insolubility in water. However, improved solubility was observed for these nucleobases affixed to ribose, as for adenosine **178**, and this allowed testing for binding affinity to receptor **3** – even then no binding affinity for **178** was seen. This is interesting as D-ribose **170** shows some minor affinity towards **3**, but substitution with nucleobases removes all activity. This is perhaps because of increased steric effects due to the presence of the pendant nucleobases but also could be due to the ribose now adopting predominantly the furanose (5-membered ring) structure, which is unlikely to bind as strongly as the pyranose (6-membered ring) form.

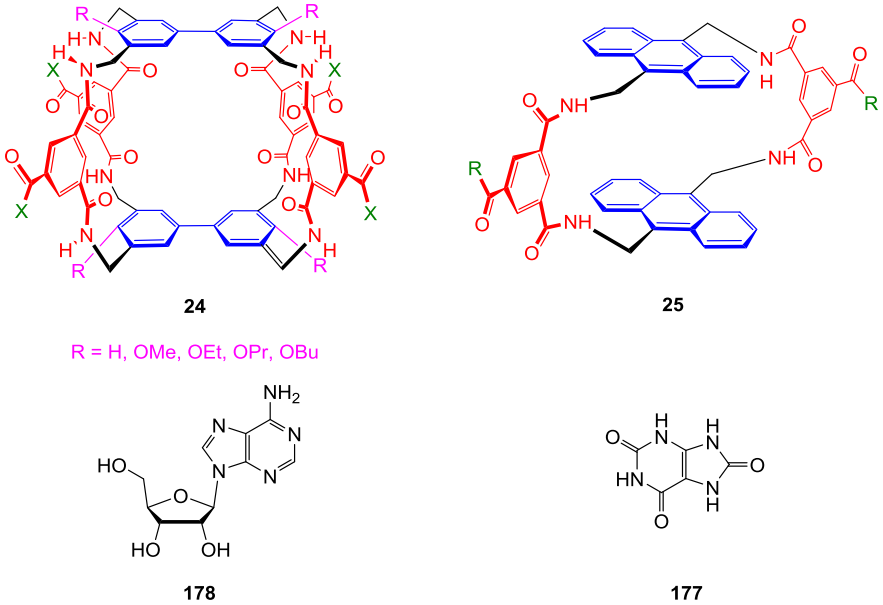
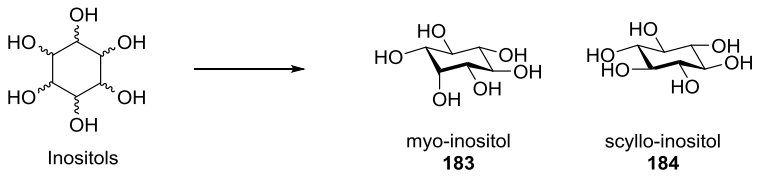


Figure 86 Previously reported synthetic lectins **24** and **25** by the Davis group, which show affinities for target substrate D-glucose **4** of K_a ~ 60 M^{-1}. However, aromatic substrates were found to bind much more strongly than D-glucose **4** to both receptors. Biphenyl receptor **24** bound adenosine **178** with K_a ~1500 M^{-1}, whereas anthracene receptor **25** bound uric acid **177** with K_a ~ 200,000 M^{-1} (with other purines bound even more strongly).[101]

The lack of affinity of **3** towards these biologically relevant substrates is important for potential medicinal applications but also because previous carbohydrate receptors based on non-covalent interactions suffered from off-target binding to such substrates, often binding them much more strongly than their intended carbohydrate guests. For example, biphenyl receptor **24** bound adenosine **178** with K_a ~1500 M^{-1} (>25 times greater than D-glucose), whereas anthracene receptor **25** bound uric acid **177** with K_a ~200,000 M^{-1} (>3500 times greater than D-glucose) and other small aromatic substrates with even higher affinities (Figure 86).[101] Therefore, the apparent lack of binding between **3** and such substrates can be considered a great advancement over previous receptor designs. This advantageous lack of affinity is believed to arise from the increased distance between the aromatic surfaces, when comparing **3** to **25** for example. The urea spacers seen for **3** hold the surfaces further apart than the previous isophthalamide spacers used for **25,** and this increased distance for **3** is believed to be too large for π-π stacking interactions between receptor and substrate, thus giving vastly reduced affinities towards aromatic substrates. The same rationale can also be applied to amino acids L-tryptophan **179** and L-phenylalanine **180**, which both feature aromatic substituents and display no observable binding to receptor **3** by ITC.

Two substrates of note are acetaminophen (paracetamol) **181** and ascorbic acid (vitamin C) **182**. Neither show any binding affinity towards receptor **3**, which is not surprising given similar substrates mentioned earlier (i.e. polyols and aromatics) also do not bind. The lack of affinity of **3** towards these substrates is very significant as they are very biologically relevant compounds, and are commonly consumed for medicinal purposes. Current commercial glucose sensing methodology based on boronic acids or proteins (such as glucose oxidase) struggles to distinguish between D-glucose **4** and these substrates (**181** and **182**) – this leads to false positives and can lead to inaccurate measurements of blood glucose concentration, something which can be detrimental to health or even fatal.[102] Given that receptor **3** does not appear bind either substrate, it possesses a potential advantage over existing glucose sensing technologies as it would not be affected by any present dosage of these common medicinal substrates.



Scheme 84 Inositols (hexahydroxy cyclohexanes) are cyclic polyols and can exist as 9 possible stereoisomers, with *myo*-inositol **183** being the most prevalent in nature. *Scyllo*-inositol **184** is the all equatorial form, and is predicted to bind to hexaurea **3** strongly given its complementary symmetry and resemblance to D-glucose **4**.

Inositols are cyclic polyols, being a derivative of cyclohexane substituted with six hydroxyl groups (Scheme 84). The most abundant in nature is *myo*-inositol **183**, which features only one axial hydroxyl with the other five being equatorial, and functions as a secondary messenger for intracellular signal transduction pathways for several biological processes.[103,104] The all-equatorial form: *scyllo*-inositol **184** is very rarely found in nature, occurring in small quantities in some tropical plants and species of shark.[105] These cyclic polyols are interesting substrates as they strongly resemble D-glucose **4**, they are in fact derived from D-glucose **4** in biology but feature more hydroxyl groups with a higher degree of symmetry.[106] *Scyllo*-inositol **184** in particular features the same C3-symmetry as receptor **3** and is therefore predicted to have a high affinity, as it can potentially take advantage of all urea groups through hydrogen bonding – something D-glucose **4** is predicted not to do by molecular modelling (Figure 60).

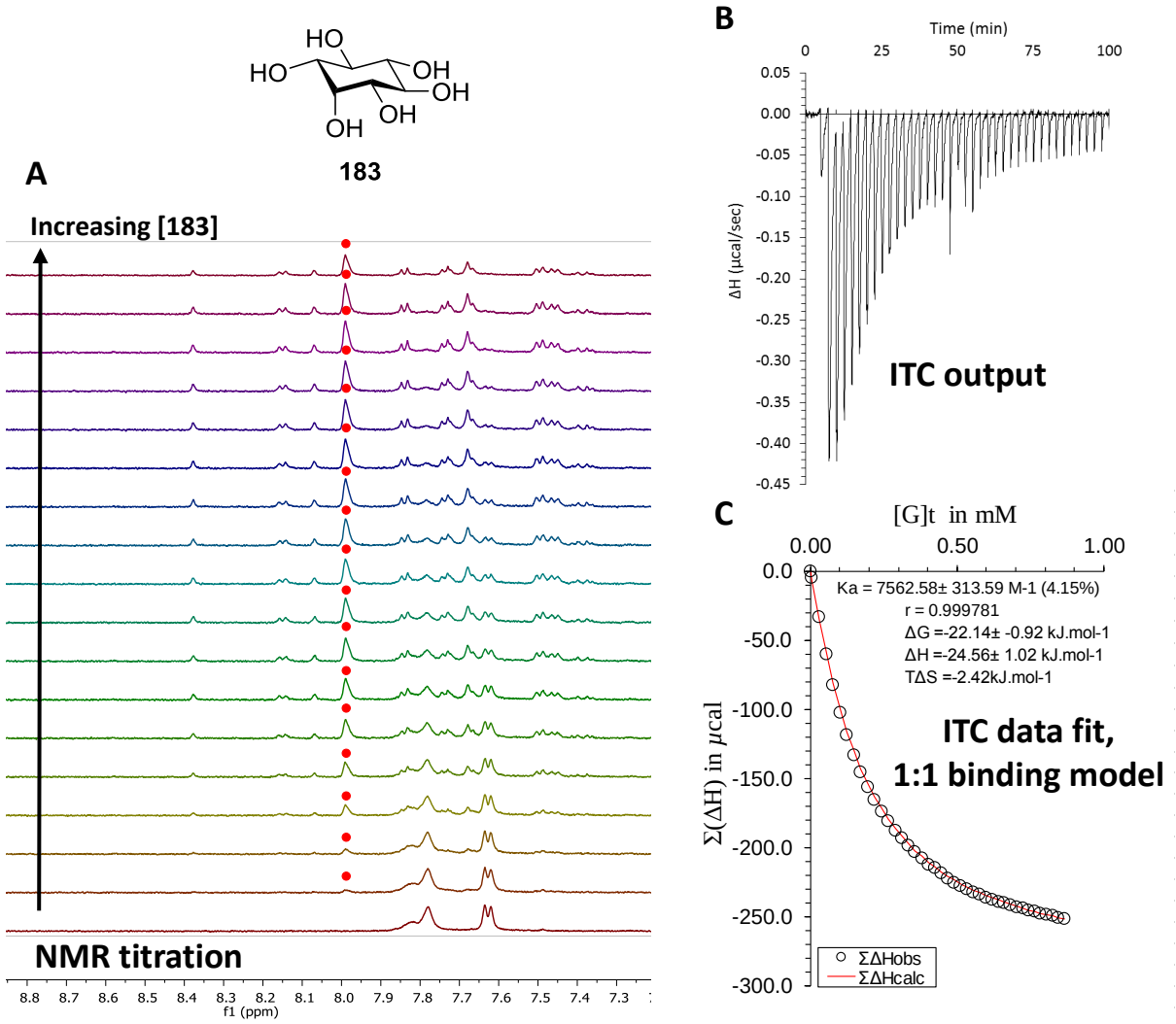


Figure 87 Binding results for *myo*-inositol **183** and receptor **3**. Partial ^{1}H NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) (A) show the NMR titration, which implies binding with slow exchange on NMR timescale. Integrations of peak at δ 7.98 ppm (denoted with •) versus region δ 8.39-7.32 ppm were plotted against guest concentration, fitting this curve to a 1:1 binding model yielded a K_a of ~7300 M^{-1}. Binding constant was also determined using ITC (B), which gave a K_a of ~7500 M^{-1} when the data was fit to a 1:1 binding model (C).

NMR binding studies of receptor **3** with *myo*-inositol **183** implied that binding was occurring with slow exchange kinetics, with signals relating to the complex evolving upon each addition of substrate. Signals related to the complex were very well defined compared to previous slow exchange spectra, and this was rationalised to be due to the substrate being one isomer (unlike D-glucose in solution) and the binding kinetics being the slowest observed thus far – with equilibrium being reached approximately 5 minutes after each addition of **183** to **3**. Integration of the spectra and fitting the data to a 1:1 binding model afforded a binding constant of K_a ~ 7300 M^{-1} (Figure 87). The same affinity was confirmed by ITC as well, with the slower kinetics of binding visualised with increased peak widths due to a slower evolution of heat upon binding. This binding affinity is relatively high considering the axial hydroxyl featured in the structure of **183**. Previously tested sugars with axial hydroxyls, such as D-

galactose **5**, displayed very low affinities (K_a <200 M^{-1}) in comparison. It is postulated that the increased affinity for *myo*-inositol **183** is due to presence of 2 more hydroxyl groups directly attached to the cyclohexane skeleton (when compared to D-galactose **5**). These must create more hydrogen bonding interactions for **183** compared to **5**, and thus result in the higher affinity observed – despite the axial hydroxyl. These increased number of intermolecular interactions as well as the increased size of the substrate (relative to most tested sugars) are likely the cause of the slower binding kinetics as well.

NMR binding studies were then performed with *scyllo*-inositol **184** and receptor **3**. Studies were initially performed using a similar concentration regime as for D-glucose **4**. However it soon became apparent, when the receptor appeared to saturate very rapidly, that the binding affinity was much higher than for D-glucose **4**. The concentration of *scyllo*-inositol **184** added was lowered accordingly and a successful titration carried out (Figure 88). As for *myo*-inositol **183**, *scyllo*-inositol also appeared to bind with slow exchange kinetics on the NMR timescale, albeit with even slower kinetics than for **183** – taking approximately 20 minutes to reach equilibrium after each addition of **184** to **3**. Nevertheless, the receptor appeared to saturate around approximately one equivalent of **3**, which already indicated very strong binding. Signals relating to the host-guest complex also appeared highly symmetrical and very well resolved, undoubtedly due to the apparent high affinity and matching symmetry of host and guest. Integration of the titration spectra and fitting the data to a 1:1 binding model gave a large binding constant of K_a ~480,000 M^{-1}, more than 26 times stronger than D-glucose **4**. Such a high affinity is to be expected, as *scyllo*-inositol **184** can be considered 'the ideal substrate' for receptor **3**.

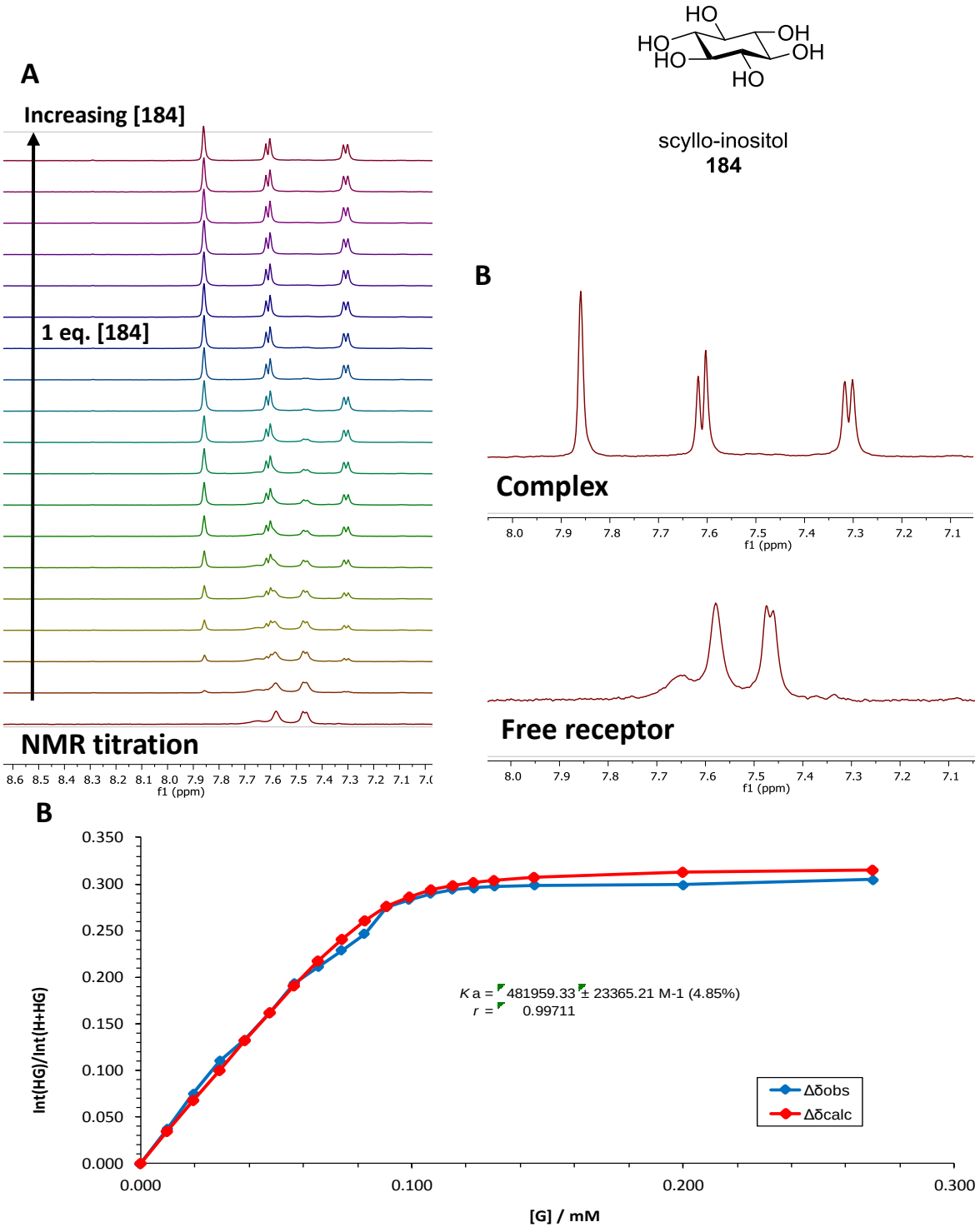


Figure 88 Binding results for *scyllo*-inositol **184** and receptor **3**. Partial ^{1}H NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) (A) show the NMR titration, which implies binding with slow exchange on NMR timescale. Signals relating to the complex (B) are very resolved and show a high degree of symmetry, presumably due to the matching symmetry of **184** and **3**. Integrations of any peak relating to the complex versus region δ 8.10-7.20 ppm were plotted against guest concentration, fitting this curve to a 1:1 binding model yielded a K_a of ~480,000 M^{-1}.

However, measuring such a high binding constant by NMR titration can be difficult due to the relatively high concentrations of host and guest required to get reliable signal intensity during the titration – especially for slow exchange binding events where signal intensity can become severely reduced if the host-guest complex is spread over a large chemical shift range. Minor errors in measured concentrations can result in very large errors for calculated binding constants, as high affinities are very sensitive to concentration of both host and guest, and these errors become more impactful with increased concentrations of host and guest used for NMR titrations. Indeed, this was observed for the titration of *scyllo*-inositol **184** with receptor **3**, where adjustment of host (or guest) concentration by 5-10% resulted in a deviation of binding constant of ~200,000 M^{-1} – with the binding curves still producing good fits with errors below 10%. Despite this, all indications suggested a binding constant of at least 10^5 M^{-1} for *scyllo*-inositol **184** and receptor **3**. However, such uncertainty in the binding affinity is not acceptable and so other methods were explored to determine accurate binding affinity, such as ITC where the use of lower concentrations of host and guest is routinely feasible.

Extraction of a binding constant from titration of **184** into **3** using ITC proved non-trivial however, due to the slow binding kinetics originally seen during the NMR titration. Applying the same ITC method used for earlier substrates resulted in one very broad evolution of heat upon addition of **184.** It wasn't until a delay of 20 minutes between additions was applied to the method that separate exotherm peaks were observed for each addition of **184**, even then the peaks were very broad and thus of low intensity – with heat being slowly evolved for over 10 minutes after addition of substrate. Not only this, but the acquisition times became highly impractical with the required 40 additions of substrate taking >13 hours to reach completion. Furthermore, the broadness of the exotherm peaks were a major issue during processing of the titration data. The software used to extract the enthalpies of each addition automatically integrates each exotherm peak to get the change in enthalpy (ΔH) at that point in the titration, but also performs a baseline correction to make integration more accurate. However, the broadness of the peaks obtained for **184** meant that the software automatically eliminated most of the heat produced upon each addition of **184** by misinterpreting this as baseline distortion, and thus smoothing out the peaks during baseline correction. This gave very inaccurate and reduced enthalpies for each addition and thus no binding constant was determinable from these data initially. However, removal of the baseline correction and manual integration of each peak did give a more consistent plot of ΔH versus concentration of **184** – albeit with a presumed large error as the limits of integration were determined visually.

This plotted data yields an 'S-curve' which is typical for binding constants greater than 10^4 M^{-1} - all previous ITC data does not exhibit an S-curve when plotted for this very reason. This S-curve arises from all substrate titrated in during the earlier additions becoming immediately complexed with the

receptor, with no free substrate remaining in solution due to the high binding constant.[100] Therefore, the same change in enthalpy (ΔH) is observed for each of these additions until a higher proportion of the receptor is bound (here until ~60% of receptor is bound) and some substrate does not participate in binding. At this point, the ΔH per addition begins to decrease to give the rest of the curve, yielding overall an S-curve. Fitting this curve to a 1:1 binding model did yield a high affinity of **3** for *Scyllo*-inositol **184** however, with a measured binding constant of K_a ~ 230,000 M^{-1}. While this measured binding affinity is approximately 50% of that when deduced by NMR titration, it was found (as mentioned earlier) that the binding constant derived from NMR was very sensitive to concentration – reducing host or guest concentration by <10% gave a binding constant similar to that obtained by ITC. It is also instructive to consider that the binding constant derived from ITC must contain some error due to the slow complexation kinetics and manual integration of the broad exotherm peaks. Therefore, these data obtained from both NMR and ITC titration would suggest that the association constant for *scyllo*-inositol **184** to receptor **3** is in the region of K_a 2-8 × 10^5 M^{-1} (upper and lower limits of error for both methods).

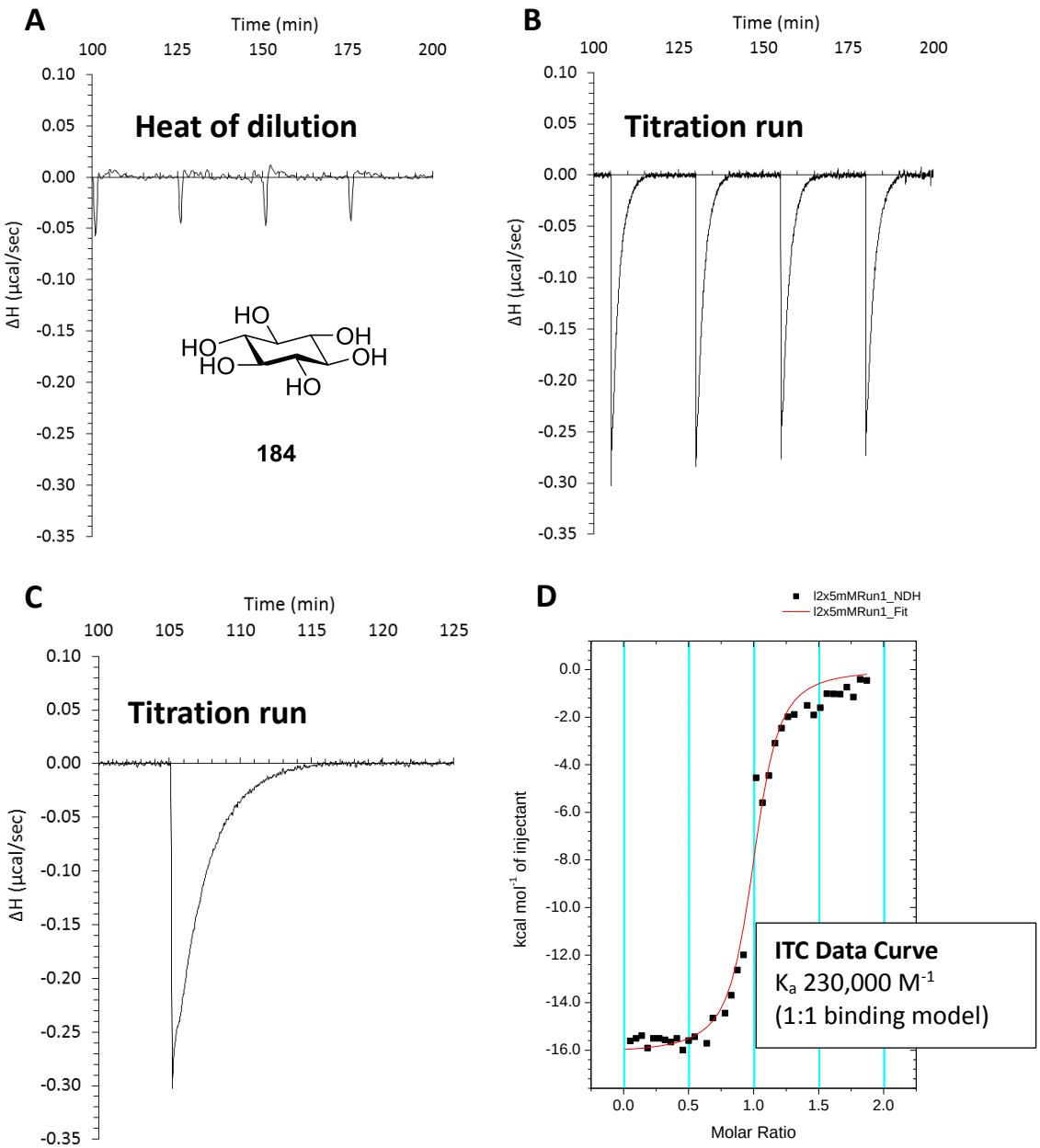


Figure 89 ITC binding results of *scyllo*-inositol **184** (2.5 mM) with receptor **3** (0.2 mM). Very large difference in exotherm observed when comparing heat of dilution (A) with titration run (B). Exotherm peaks (C) exhibit slow evolution of heat (>10 mins), which are evidence of slow complexation kinetics. Plotting this data as change in enthalpy (ΔH) versus concentration (D) and fitting to 1:1 binding model gives association constant of K_a ~230,000 M^{-1}.

When determination of a large binding constant is impractical due to minor errors in concentration of host or guest, alternative methods to accurately determine the binding constant must be considered. The easiest solution would be to lower both host and guest concentrations used, in order to minimise deviations in concentration upon creation of titration solutions. However, as described before, this can

prove impractical when concentration is directly responsible for signal intensity for methods such as NMR. An alternative approach would be to perform a competitive titration, in this case by NMR.[16] This procedure involves pre-saturation of the host receptor with a guest of a known (usually lower) binding affinity, with the higher affinity guest then titrated into this already saturated system. The presence of the weaker binding guest effectively lowers the binding constant of the stronger binding guest in this system, due to the equilibrium now also being governed by decomplexation of the weaker binding guest – lowering the Gibb's free energy of binding (ΔG) and thus association constant (K_a). This competitive titration system is governed in part by the following equilibria (where A is the weaker binding guest and B is the stronger binding guest, i.e. *scyllo*-inositol):

$$A + H \underset{K_{H+A}}{\overset{K_{HA}}{\rightleftharpoons}} HA \quad \text{(Eqn. 16)} \qquad B + H \underset{K_{H+B}}{\overset{K_{HB}}{\rightleftharpoons}} HB \quad \text{(Eqn. 17)}$$

A = weaker binding guest B = stronger binding guest

Equations that govern complexation in these equilibria are as follows:

$$K_{HA} = \frac{[HA]}{[H][A]} \quad \text{(Eqn. 18)} \qquad K_{HB} = \frac{[HB]}{[H][B]} \quad \text{(Eqn. 19)}$$

Therefore, in the proposed competitive titration, the complexation of the stronger binding guest can be represented as:

$$\frac{K_{HB}}{K_{HA}} = \frac{[HB][A]}{[HA][B]} \quad \text{(Eqn. 20)}$$

Fortunately, if both guests (A and B) exhibit binding kinetics that are slow on the NMR timescale, most of these parameters can be calculated from the integration of a single spectrum. The ratio of [HB]/[HA] can be determined directly from integration of signals relating to both complexes, assuming an appropriate ratio of each complex had been reached upon equilibrium (i.e. 1:1 [HA]:[HB]). Values for [A] and [B] are simply calculated from the following:

$$[A] = [A]_{total} - [HA] \quad \text{(Eqn. 21)} \qquad [B] = [B]_{total} - [HB] \quad \text{(Eqn. 22)}$$

The association constant of guest A (K_{HA}) is calculated from previous titration data, and is thus known. Despite the opportunity to calculate K_{HB} from a single slow exchange spectrum, it was initially decided to perform a titration to maximise the amount of calculated binding constants from each titration spectra, to obtain the most accurate value for K_{HB}.

The weaker binding guest was initially chosen to be *myo*-inositol **183**, as it bound to receptor **3** in slow exchange on the NMR timescale with a weaker binding constant (~2 orders of magnitude lower, K_a ~

7500 M^{-1}), but also because both complexes of receptor with **183** or **184** featured distinct NMR signals that did not overlap – allowing more accurate determination of [HB]/[HA]. Initially a ratio of 20:1 (**183:184**) was used, but this was then repeated with a larger ratio of 100:1 (**183:184**) to allow a more gradual titration and thus determine a more accurate value for K_{HB}. However, it was discovered upon processing of the titration data that an anomalous increase in the ratio of [HB]/[HA] had occurred part way through the titration (Figure 90) – even though a constant amount of *scyllo*-inositol **184** was being added before and after this anomalous result. The other points in the titration appeared very consistent, with a linear increase in the ratio of [HB]/[HA] after addition of scyllo-inositol **184**. The only variable was equilibration time, where the anomalous was result was left for a much longer time (90 minutes vs 20 minutes for the other additions) before acquisition of the spectra. This would imply that equilibration between the *myo* **183** and *scyllo*-inositol **184** complexes was very slow after addition of *scyllo*-inositol **184**, and the 20 minutes equilibration period was insufficient. Indeed, following test studies aimed at generating a rough 1:1 ratio of *myo* **183** and *scyllo*-inositol **184** complexes showed that equilibrium had not been reached after 1 day. It was then hypothesised that *myo*-inositol **183** may be unsuitable as a competitor, as it is the substrate with the second slowest binding kinetics tested. Given that the slowest binding kinetics are observed for *scyllo*-inositol **184** itself, these slow binding kinetics may compound one another and result in an incredibly slow equilibrium, especially if the rate of decomplexation of *myo*-inositol **183** or the rate of binding of *scyllo*-inositol **184** are rate limiting

A new competitor was needed and it was decided that methyl β-D-glucoside **14** would be suitable for several reasons: it exhibits faster binding kinetics than *myo*-inositol **184** while still operating with slow exchange kinetics relative to the NMR timescale (so the previous methodology used to calculate K_a can still be used); both receptor and guest complexes (for **14** and **184**) do not have overlapping NMR signals and so can be accurately integrated; and it has essentially the same binding constant as for **183** so previous competitive titration data can be used to predict experimental parameters for this new competitive study. A slightly different experiment was proposed than the previous competitive titrations however, with two simultaneous studies planned. One experiment would feature addition of *scyllo*-inositol **184** to receptor **3** saturated with methyl β-D-glucoside **14** (equivalents ratio of 2:1:400, **184:3:14**). At the same time, the opposite addition would be performed where **14** would be added to **3** saturated with **184** (same equivalents ratio). NMR spectra would then be recorded until equilibrium had been reached – this would now be obvious as the measured ratio of both complexes (i.e. [HB]/[HA]) would be equal for both samples.

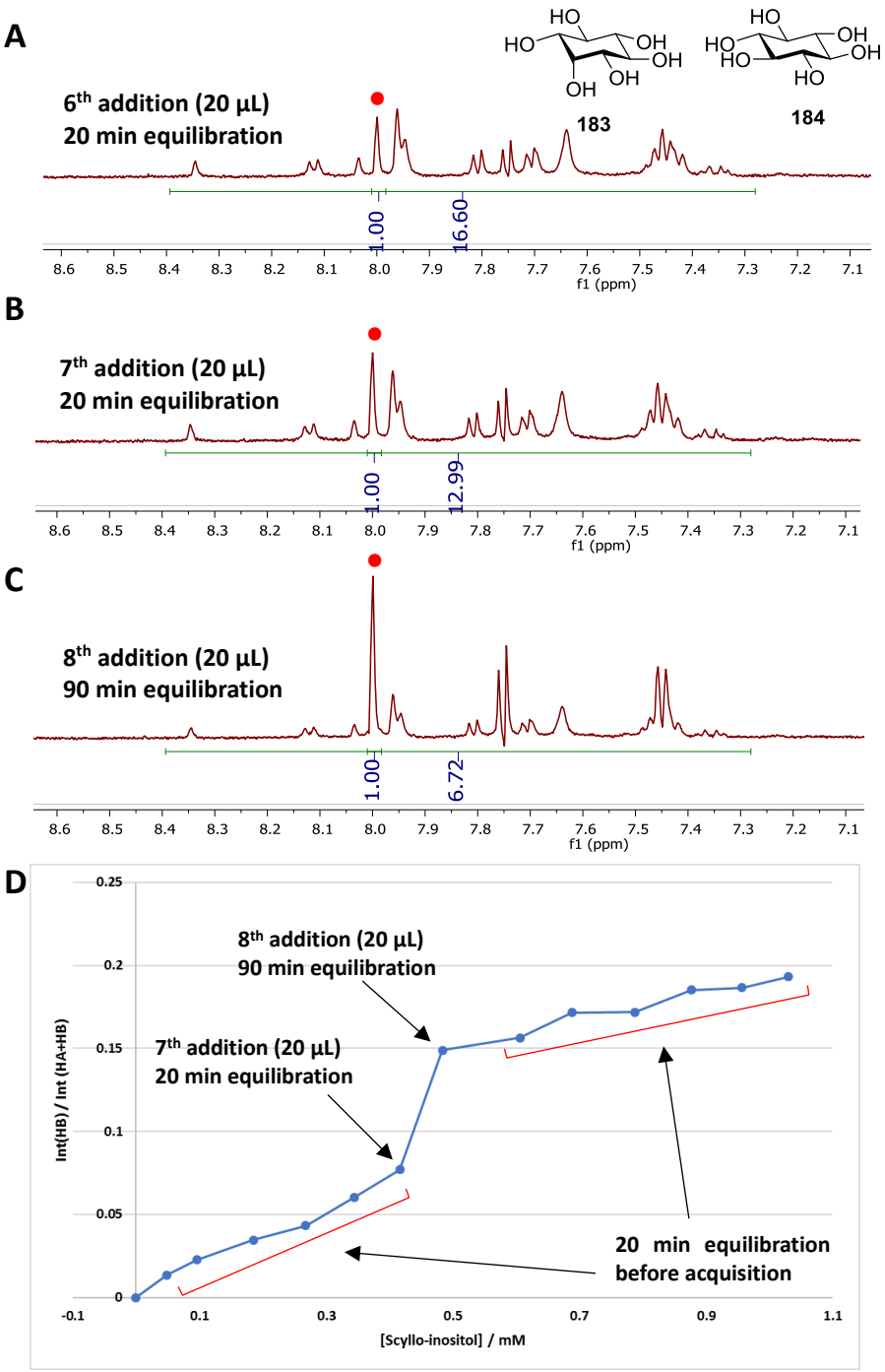
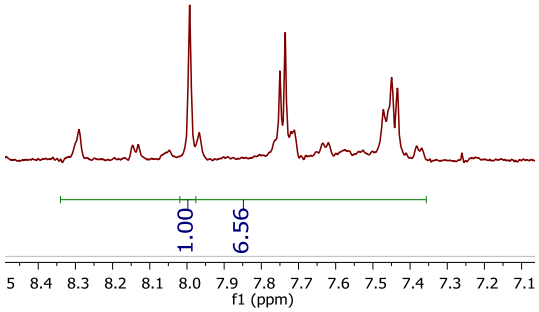


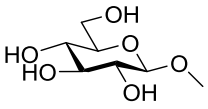
Figure 90 Partial ^{1}H NMR of titration spectra (600 MHz, D_2O, pH 7.4, 298 K) from addition of *scyllo*-inositol **184** (8.5 mM) to receptor **3** (0.2 mM) saturated with *myo*-inositol **183** (40 mM). Ratio of complexes ([HB]/[HA]) determined by integration of peak at ~8 ppm (denoted with •) versus region 7.3-8.4 ppm. Earlier additions during titration (A and B) featured an equilibration time of 20 minutes before acquisition and a consistent evolution of *scyllo*-inositol **184** and receptor **3** complex was observed. Subsequent addition (same volume of guest added, spectrum C) was equilibrated for 90 minutes and a large increase in *scyllo*-inositol complex was observed (see difference in integrations and graph D). Additions after this point were then equilibrated for 20 minutes prior to acquisition and linear trend continues (graph D). These data imply that very slow conversion to the *scyllo*-inositol complex is taking place, and therefore that almost all integrations to determine [HB]/[HA] during this experiment are incorrect.

However, it became apparent that the rate of equilibration was still incredibly slow, with the amount of *scyllo*-inositol **184** and receptor **3** complex still increasing after 18 hours. Both samples were then left for approximately 2 weeks, with the intention that both would have certainly reached equilibrium by then. Indeed, integration of both samples after this period indicated that approximately 30% consisted of *scyllo*-inositol **184** and receptor **3** complex, with the other 70% of receptor **3** bound to the competitor methyl β-D-glucoside **14**. Analysis of the integrals over time and extrapolation of this data suggested an equilibration time of 4-5 days, which is incredibly slow when compared to equilibration times seen for free receptor and other guests. It was noted that the predicted equilibration time was 1.5 times slower for *scyllo*-inositol **184** displacing methyl β-D-glucoside **14** from complexation with receptor **3** – this would imply that either decomplexation of **14** or complexation of **184** is rate limiting. Given that **184** exhibits the slowest complexation kinetics with free receptor, it is then presumed that the binding of **184** to the receptor is rate limiting – possibly due to its larger size when compared to **14**. For the reverse system, where **14** is added to **184** complexed to **3**, the slow kinetics are presumed to be due to the high binding affinity of **183** reducing the rate of dissociation.

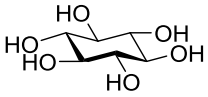
Applying the integrated ratio of complexes to the previously outlined methods for calculating binding constants for a competitive binding system afforded an association constant of K_a ~800,000 M^{-1} for *scyllo*-inositol **184** and receptor **3** (~45 times stronger than D-glucose). This value is towards the upper end of error for the previously determined binding constants, but those methods (1:1 NMR and ITC titrations) potentially carried a lot of error due to the issues outlined earlier. Future work could employ different competitors and acquire a larger sample size of data to derive the binding constant from – something that would reduce the error in this value. A study at increased temperatures would result in faster kinetics and thus equilibrations times, allowing faster acquisition of data. However, a study of such scale was not feasible due to insufficient time. It is worth mentioning that such a high binding constant for a neutral polar molecule (or “carbohydrate -like” molecule) in water with non-covalent interactions is largely unprecedented, and represents a significant advancement in the area of molecular recognition in water.

A





[A] 14



[B] 184

B

$$[HB] = \frac{Int(HB)}{Int(HA + HB)} \times [H]total$$

$$[HA] = [H]total - [HB]$$

$$[A] = [A]_{total} - [HA]$$

$$[B] = [B]_{total} - [HB]$$

$$\frac{K_{HB}}{K_{HA}} = \frac{[HB][A]}{[HA][B]}$$

C

Raw Integration from NMR	6.56	
[HA]/[HB]	2.19	
Percentage of [HB]	31.4%	
[H]total	135	µM
[HB]	42.4	µM
[HA]	92.8	µM
[A]total	54	mM
[A]	53.907	mM
[B]total	270	µM
[B]	228	µM
K_{HA}	7500	M^{-1}
K_{HB}	**8.11E+05**	M^{-1}

Figure 91 Competitive binding studies of *scyllo*-inositol **184** and methyl β-D-glucoside **14** with receptor **13** (equivalents ratio of 2:400:1, **184**:**14**:**3**). Integration of NMR spectrum (600 MHz, D_2O, pH 7.4, 298 K) (A) once system reaches equilibrium gives ratio of complexes (i.e. [HB]/[HA]). Applying this ratio to competitive binding equations (B) allows calculation of binding constant (C) for *scyllo*-inositol **184** to receptor **3**. Calculated binding affinity is K_a ~811,000 M^{-1} using this method.

Despite this very high binding affinity of *scyllo*-inositol **184** for receptor **3**, *scyllo*-inositol **184** is not commonly found in nature.[105] Thus, D-glucose **4** remains the strongest binding molecule that could be considered 'biologically relevant', with still excellent selectivity over other biologically relevant substrates. To further ascertain how robust the binding and selectivity of D-glucose can be in a more representative biological environment, binding of D-glucose in complex mixtures of organic molecules, salts and variations in pH was investigated. Binding studies to D-glucose by ITC were then undertaken in standard phosphate buffered saline (PBS, for composition see experimental) at pH 6 ,7 and 8. A large exotherm was observed for each titration and fitting the data to a 1:1 binding model gave K_a ~ 17,300-18,300 M^{-1}, essentially the same affinity as in water. PBS contains significant quantities of sodium and potassium chloride, and it would appear that such salts do not have an effect on binding affinity.

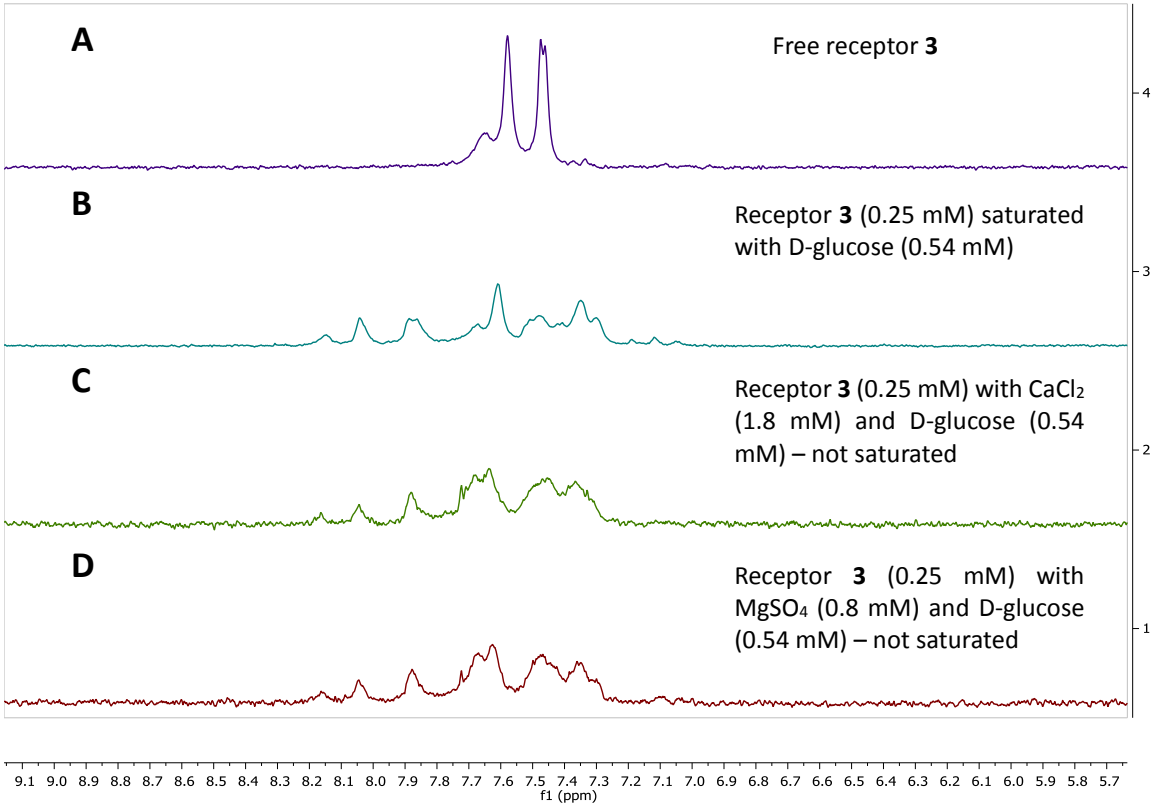
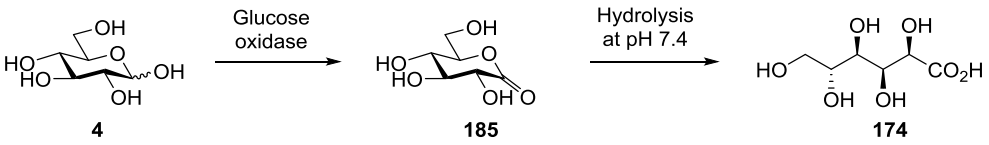


Figure 92 Control NMR studies (600 MHz, D_2O, pH 7.4, 298 K) investigating the effects of calcium and magnesium cations on binding of D-glucose **4** to receptor **3**. Free receptor **3** (A) at 0.25 mM is saturated with 0.54 mM of D-glucose **4** (B). Addition of $CaCl_2$ (1.8 mM) or $MgSO_4$ (0.8 mM) appears to suppress binding (C and D), with the same amount of D-glucose **4** (0.54 mM) not saturating receptor **3**.

Binding of **3** to D-glucose **4** in two types of cell culture media was then tested. Dulbecco's Modified Eagle Medium (DMEM) and Leibovitz's L-15 (L-15) culture media were chosen as neither contained any D-glucose **4** but both contained significant quantities of amino acids, vitamins, sugars and inorganic salts (for formulations see Chapter 7, Table 8 and Table 9) – a very demanding environment for the receptor to still function in.[107,108] ITC binding studies with D-glucose **4** in both media did show that receptor **3** still bound D-glucose, albeit with a lowered affinity of K_a ~5200-5600 M^{-1}. Given that receptor **3** showed no binding to several amino acids, vitamins and very weak binding to sugars in isolated studies, it was rationalised that the reduced affinity might be due to the high concentrations of inorganic salts present in both media. A control medium was then created which was comprised of only inorganic salts at the same concentrations found in the DMEM cell culture medium (for formulation see Chapter 7, Table 7). ITC titrations of D-glucose **4** and **3** in this medium gave effectively the same results as for the DMEM medium, with a measured binding constant of K_a ~5200 M^{-1}. This would confirm that the salts (or one specific salt) is the cause of the lower binding constant. Further control experiments concluded that the reduced binding was due to calcium chloride and magnesium sulfate salts. These are present at millimolar concentrations in the culture media, and partial binding

studies indicated that higher concentrations of D-glucose were needed to saturate receptor **3** in their presence (relative to in just pure water). It is as of yet not fully known by what mechanism these cations suppress the binding affinity, as control experiments and previous literature suggest the metal cations do not perturb D-glucose itself.[109] A minor change in chemical shift for NMR signals relating to the dendritic solubilising groups was observed and it is well known that these divalent cations (Ca^{2+} and Mg^{2+}) can associate with up to two carboxylates (i.e. $Mg^{+}(RCOO^{-})_2$).[110] Given that the solubilising groups are comprised of carboxylates (with receptor **3** featuring 27 carboxylates in total), it is theorised that these metal cations could effectively 'cross-link' the solubilising group network and hinder entry to the cavity. They could even bind to the linker carbonyls and distort the shape of the receptor, which would undoubtedly affect the binding affinity to D-glucose.

Although these cations (Ca^{2+} and Mg^{2+}) do appear to diminish the binding affinity for D-glucose, they do not occur in such high concentrations in the human body or blood, and would have a reduced effect if applied to a system more representative of these environments. As such, the binding affinity of D-glucose in human serum was investigated. Human serum is comprised of all components of whole blood minus blood cells and clotting factors – so consists of proteins, antibodies, hormones, salts and many small molecules, and is commonly used in both healthcare and research environments as a model system for studying blood.[111] For the purposes of investigating the binding of **3** to D-glucose **4** in this medium, the first step is to remove the large amount endogenous D-glucose. Concentrations of D-glucose in human serum were measured using a YSI 2300 STAT Plus D-glucose and L-lactate analyser, with the initial concentration in human serum measured as 6.2 mM.[112] The D-glucose was removed from the serum using enzymes glucose oxidase and catalase, which firstly oxidise the D-glucose to D-glucono-δ-lactone **185** (Scheme 85) which then hydrolyses in the medium to D-gluconic acid **174**, which is known not to bind to receptor **3** through previous binding studies (Figure 84).[113] Analysis of the medium after oxidation (using the YSI 2300) suggested oxidation was successful, with the concentration of D-glucose measured as 0-0.005 mM.



Scheme 85 Oxidation of D-glucose **4** to D-glucono-δ-lactone **185** with glucose oxidase. Hydrolysis of **185** at biological pH (7.4) yields acyclic polyol gluconic acid **174**, which is known to not bind to receptor **3** through previous binding studies.

Removal of high molecular weight components (such as antibodies, and large proteins) is necessary as ITC is very sensitive to anything that produces heat, such as precipitation and secondary binding events like aggregation, and phenomena like these can either be misinterpreted as or mask binding events and thus the influence results obtained.[114] Filtration of the serum post oxidation of D-glucose through a membrane (10,000 MWCO) to remove the enzymes and high molecular weight components gave a solution which retains all the small proteins/peptides, molecules and salts that might bind to receptor **3**. It was this medium that was used for the ITC binding studies of D-glucose and receptor **3** in 'human serum'.

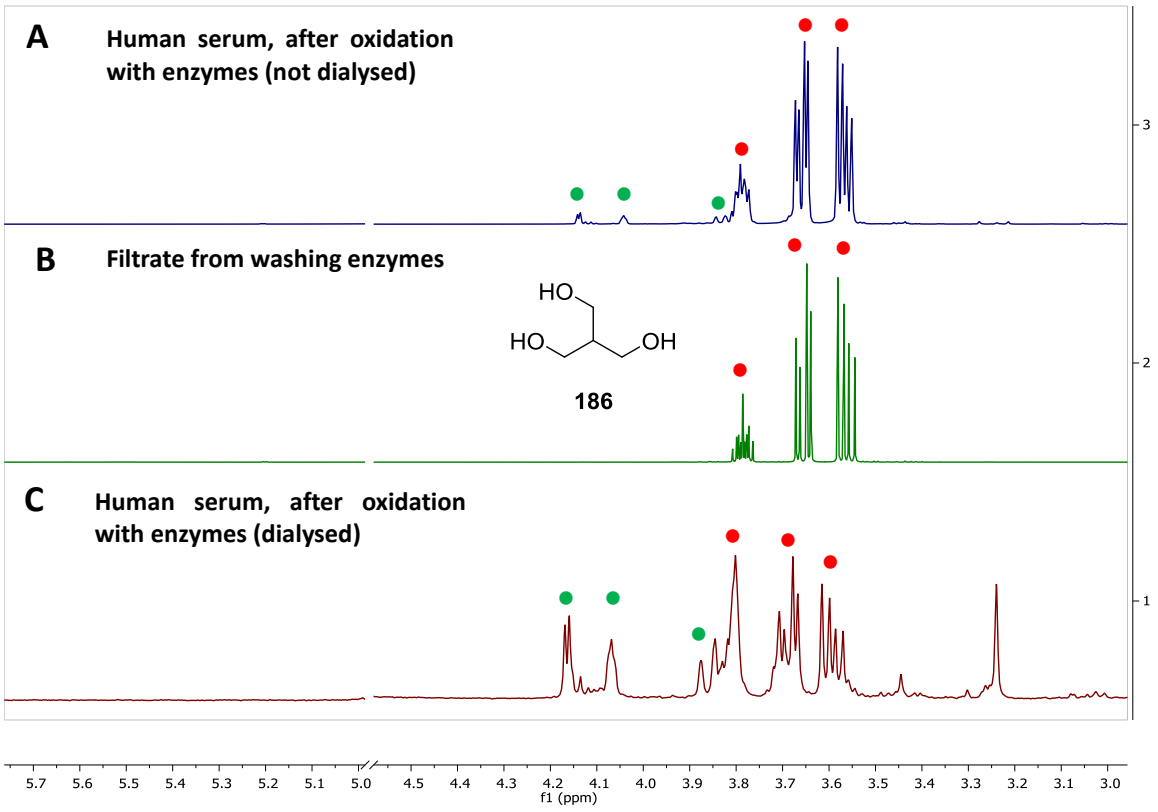


Figure 93 Partial ^{1}H NMR spectra (600 MHz, D_2O, pH 7.4, 298 K) of human serum after oxidation of D-glucose, with glucose oxidase and catalase, which shows a significant quantity of glycerol **186** present (A, denoted with •). Filtration of enzymes and analysis of filtrate shows only glycerol (B), suggesting glycerol originates from enzymes used for oxidation of D-glucose. Dialysis of enzymes in pure water and repeating oxidation procedure shows vastly reduced amounts of glycerol **186** in human serum, with the signals for D-gluconic acid **174** (denoted with •) now much increased in intensity (relative to those in spectrum A).

Initial ITC binding studies of D-glucose **4** to **3** in human serum, gave an association constant of K_a ~2500 M^{-1}, even lower than observed for D-glucose in cell culture media. It was difficult to rationalise how the binding constant could be even lower than the cell culture media given the limited data provided by ITC, so several control experiments were performed. The controls were initially to ascertain whether the enzymes used to 'remove' D-glucose from the human serum (glucose oxidase and catalase) were

the cause of the lower affinity, either by impurities carried through, or by incomplete oxidation of all D-glucose and then failure by the YSI 2300 to detect any residual D-glucose. A control sample of 6 mM D-glucose in water was prepared and exposed to the same enzymes (glucose oxidase and catalase) as the human serum. Filtration of this medium through a membrane (10,000 MWCO) removed the enzymes and an ITC titration with D-glucose in this medium gave a similarly reduced affinity as for the human serum – association constant for **3** and D-glucose **4** of K_a ~4000 M^{-1}. Analysis by ^1H NMR of this control medium after filtration revealed that all of the D-glucose had indeed been converted to D-gluconic acid **174**, so the reduced binding affinity cannot be due to residual D-glucose. What was observed however, was a significant amount of glycerol **186** – which was present in amounts many times greater than the D-gluconic acid **174**. Freeze drying the human serum used for the ITC titration (after removal of D-glucose and enzymes) and reconstitution in D_2O allowed analysis of the human serum by NMR, which also revealed that significant quantities of glycerol **186** were present. Based on evidence from the 'control medium', and the fact that D-gluconic acid **174** shows no evidence of binding affinity by NMR or ITC titrations in isolated studies, the reduced binding must be caused by the presence of large quantities of glycerol **186**. It is rationalised that if **186** were present in such significant quantities, it could effectively alter the solvent composition of the medium which would undoubtedly affect binding to the receptor – as solvent effects are arguably one of the biggest factors that influence how favourable binding interactions can be.[115] Future studies would need to be performed to fully understand the effects of solvent composition (including glycerol) on binding affinity.

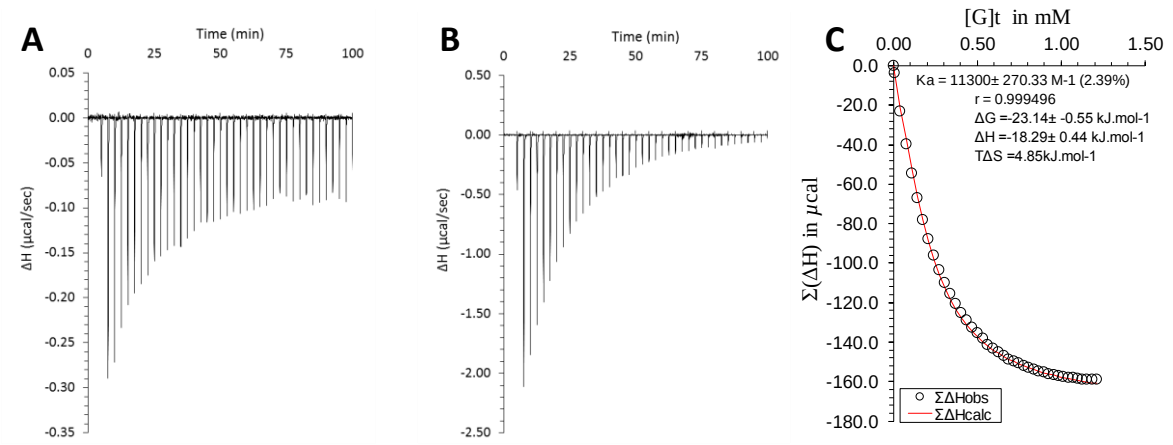


Figure 94 ITC binding results for D-glucose **4** and receptor **3** in human serum. ITC titration without dialysing enzymes, and thus presence of large amounts of glycerol **186**, results in weaker binding and lower heat output (A). Removal of glycerol **186** from enzymes by dialysis, results in larger heat output upon addition of D-glucose **4** to **3** (B, note y-axis values) and thus stronger binding affinity (C) – measured K_a ~11,300 M^{-1}.

Analysis of the human serum by ^{1}H NMR, before use of the enzymes to oxidise D-glucose, revealed clear signals for D-glucose but no signals relating to glycerol **186**. It was therefore assumed that the glycerol **186** must be introduced *via* the enzymes. Filtration (10,000 MWCO membrane) and washing of the enzymes with water confirmed this, as the filtrate contained predominantly glycerol **186** when analysed by ^{1}H NMR. Present literature indicates that glycerol **186** is added to enzymes before they are lyophilised, to preserve the protein structure upon freezing before sublimation.[116] Unfortunately, the presence of this glycerol **186** was not indicated by the manufacturer. The procedure to remove the D-glucose from the human serum was thus modified, with the glucose oxidase and catalase being dialysed (10,000 MWCO) in pure water to remove the glycerol **186** from the enzymes. Repeating the oxidation of D-glucose in the human serum with these dialysed enzymes, and filtration through a membrane (10,000 MWCO), gave a clear solution. Analysis of this medium by ^{1}H NMR revealed the presence of D-gluconic acid **174** and a very complex mixture of other molecules – but most importantly the amount of glycerol **184** was now severely reduced (presumably now at standard biological concentration). Gratifyingly, an ITC titration in this medium (with vastly reduced glycerol **186**) gave a good affinity for D-glucose **4** and **3**, with a measured K_a ~11,300 M^{-1} – roughly a 30% reduction in activity compared to pure water. The reduced affinity when compared to in pure water is presumably due to many weakly binding molecules needing to be displaced by D-glucose from the receptor **3**, or presence of magnesium/calcium cations as discussed earlier. However, the relatively high affinity would indeed suggest that the receptor is still proving to be very selective towards D-glucose, even in such a complex mixture of molecules and salts as human serum. Such affinity and excellent selectivity would make receptor **3** a promising candidate towards recognition of D-glucose in a biological context.

Despite sufficient binding data being acquired from NMR and ITC titrations, other methods were also pursued to quantify binding affinity. Spectroscopic techniques like fluorescence are of particular interest, as if a quantifiable change in fluorescence emission upon binding can be measured, this would be an early proof of principle in the application of receptor **3** as a D-glucose sensor.[15] A fluorescence excitation spectrum of receptor **3** revealed an absorption maximum at ~270 nm, which if then used as an excitation wavelength for **3** produced an emission maximum at 370 nm – a large stokes shift of ~100 nm. It was this excitation wavelength of 270 nm that was used during the fluorescence titration of D-glucose **4** and **3**, while monitoring intensity at emission wavelength of 370 nm. However, it was found that only a very small increase in emission intensity was observed upon addition of D-glucose to **3**. Such a small increase in intensity introduced large errors when attempting to plot the data and fit the curve to a 1:1 binding model, and so no binding constant was determinable from the fluorescence titration data. Previous synthetic lectin systems such as anthracene receptors **25** and **1** (Chapter 4), provided a large increase in fluorescence emission upon binding carbohydrates.[35] This was rationalised

to be due to a change in conformation upon binding, which reduced the fluorescence quenching of the anthracene surfaces. Other synthetic lectin systems that were more rigid in conformation, such as **32** (Chapter 1, Figure 18), displayed very little to no change in fluorescence upon binding – presumably due to this lack of conformational change upon binding.[39] It is therefore hypothesised that something similar is occurring for receptor **3** as well, where the free receptor is sufficiently rigid in conformation so that there is little change upon binding of D-glucose.

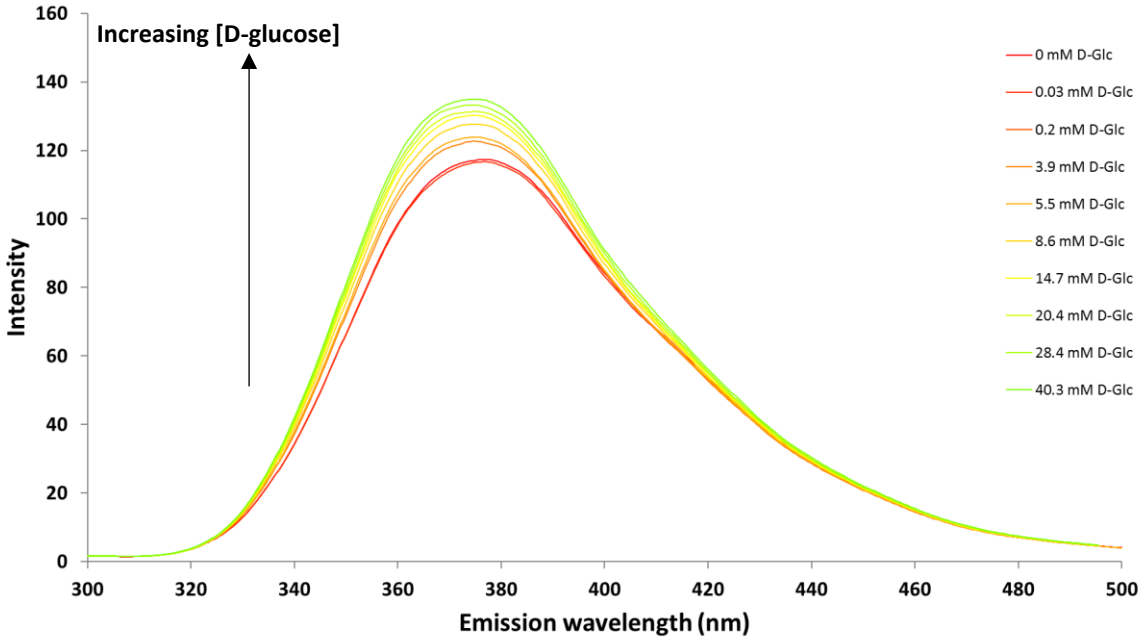


Figure 95 Fluorescence titration results from addition of D-glucose **4** to receptor **3** in water, excitation wavelength of 270 nm and monitoring at wavelength of 370 nm. Very small increase in fluorescence emission upon increasing concentration of D-glucose **4**. This introduced a large degree of error when fitting data to binding model and thus no binding constant was determined from these data.

With no determination of binding constant from fluorescence titration, other spectroscopic methods were considered, such as circular dichroism (CD). This technique has been applied to studying the complexation of molecules before, where the complexation of achiral (with a chromophore) and chiral (no chromophore) binding partners produced a CD absorption band. Such a phenomenon is known as induced circular dichroism (ICD), an example of which is the complexation of β-cyclodextrin (chiral, no chromophore) with cyclohexanone (achiral, chromophore) – which afforded an ICD band due to the n→π* transition of the carbonyl group of the cyclohexanone.[73,117] This induction of a CD signal occurs due to the chiral partner distorting the conformation of the achiral UV-active partner upon complexation, this destruction of symmetry results in a UV-active chiral complex that produces a CD absorption band. This process was predicted to translate well to the binding of receptor **3** (achiral, chromophore) and D-glucose **4** (chiral, no chromophore), with an expected ICD upon complexation.

Indeed, such a result was obtained, whereby titration of D-glucose **4** to receptor **3** in water resulted in a CD absorption of negative intensity (Figure 96). Fitting this data to a 1:1 binding model yielded a K_a of ~17,200 M^{-1}, essentially the same as for other methods (NMR, ITC). Titration of L-glucose (opposite enantiomer to naturally occurring D-glucose) into receptor **3** gave essentially the mirror image of the CD spectrum for D-glucose, with a CD absorption of positive intensity. Processing of this data also afforded the same binding constant of K_a ~17,200 M^{-1}, which is expected due to the achiral nature of receptor **3**. Distinguishing these enantiomeric complexes by NMR and ITC would not have been possible. The binding affinity was also confirmed using CD titration for carbohydrate substrates which exhibited intermediate exchange kinetics relative to the NMR timescale, which meant no binding constant was determined by NMR titration. Affinities determined using CD for D -xylose **149** (K_a ~ 5000 M^{-1}), D-glucuronic acid **168** (K_a ~ 4400 M^{-1}) and 2-deoxy-D-glucose **169** (K_a ~ 680 M^{-1}) were all very comparable to the equivalent ITC data, further corroborating these results.

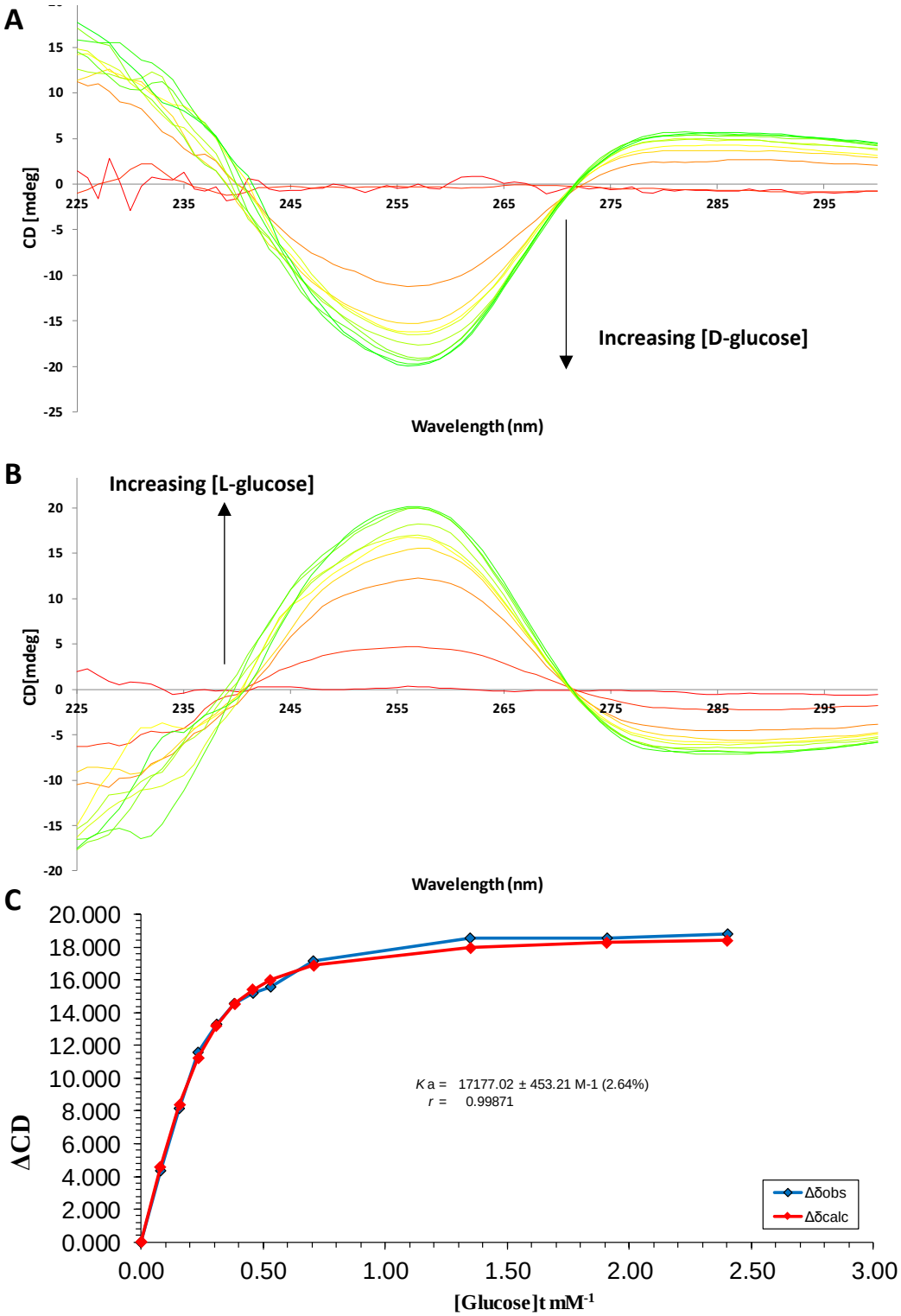
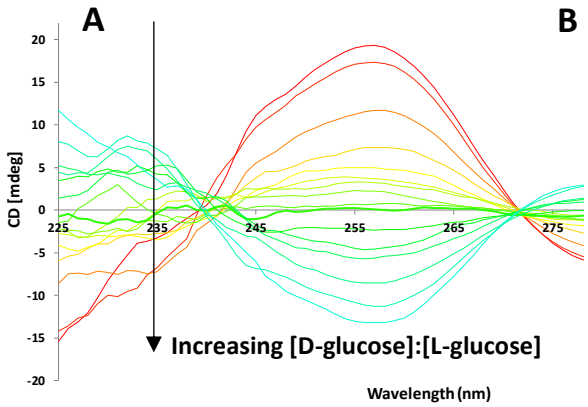


Figure 96 ICD binding results for D-glucose (A) and L-glucose (B) with receptor **102**. Both chiral guests induce a CD signal upon addition to **3**, with each enantiomer giving the opposite absorption intensity. Fitting this data to a 1:1 binding curve gave an association constant similar to other methods (C), K_a ~17,200 M^{-1}. The binding curve shown is for L-glucose, but was essentially the same for D-glucose.

A novel experiment was then proposed, where receptor **3** was pre-saturated with L-glucose of known concentration (here 2.4 mM of L-glucose) in water. A mixture of D-glucose, with **3** and L-glucose (2.4 mM) added to keep their concentrations constant, was to be titrated into this medium and it was predicted that the CD signal would equal zero when an equal racemic mixture of complexes had been reached (i.e. 1:1, L-glucose:D-glucose complexes). Most pleasingly, such a result was obtained, where the CD signal intensity reached zero upon the addition of D-glucose that resulted in an equal concentration of D and L-glucose in solution – and thus an equal concentration D and L-glucose complexes (Figure 97).



[D-Glc] mM	CD [mdeg]	[DGlc]:[LGlc]
1.86	2.78	0.77
2.08	1.68	0.86
2.29	0.71	0.96
2.40	-0.02	1.00
2.93	-2.23	1.22
3.44	-4.07	1.44
3.95	-5.18	1.64

Figure 97 ICD results from titration of D-glucose into receptor **3** pre-saturated with L-glucose (2.4 mM). **3** and L-glucose (2.4 mM) were added to titrant to keep their concentrations constant. Spectra decrease in intensity upon increased additions of D-glucose (A). CD signal reaches zero intensity when concentrations of D and L-glucose are equal (B), and thus concentrations of D and L-glucose complexes are also equal.

One application of this ability to manipulate the CD signal depending on the ratio of D and L-glucose bound to receptor **3**, is quantification of D-glucose in various media such as human serum. As titration of L-glucose into human serum (with receptor **3** present) would give zero CD signal when the ratio of D and L-glucose complexes is equal – the concentration of L-glucose added at his point would therefore be equal to the concentration of D-glucose in the human serum. However, before attempting this, analysis of the human serum medium itself by CD was undertaken to further probe how effective removal of the D-glucose (Scheme 85) by the glucose oxidase and catalase was.

Analysis of unfiltered human serum by CD resulted in observation of large absorption artefacts (Figure 98), presumably due to the presence of high molecular weight biological structures (proteins and antibodies etc.) or scattering of light by insoluble particles. Filtration of the human serum through a membrane (10,000 MWCO) removed these artefacts, with no observed CD absorption. It presumed that any specific UV-active chiral molecules (e.g. amino acids such as L-tryptophan) are present at low enough concentration to not produce an observable CD absorption at the wavelengths monitored

(225-325 nm). Addition of receptor **3** to this medium resulted in a distinct CD absorption, due to the binding of **3** to D-glucose already present in the human serum. Removal of D-glucose from the human serum using glucose oxidase and catalase, and filtering through a membrane (10,000 MWCO), again showed no CD absorption – even upon addition of receptor **3**. This would indeed imply that removal of D-glucose using the enzymes is effective, as earlier titrations in water by CD indicate D-glucose concentrations as low as 0.07 mM give a very noticeable CD absorption. A titration of D-glucose into receptor **3**, in human serum (with D-glucose removed), was then performed and gave a binding constant of K_a ~10,300 M^{-1} – corroborating the result obtained by ITC (described earlier in this chapter).

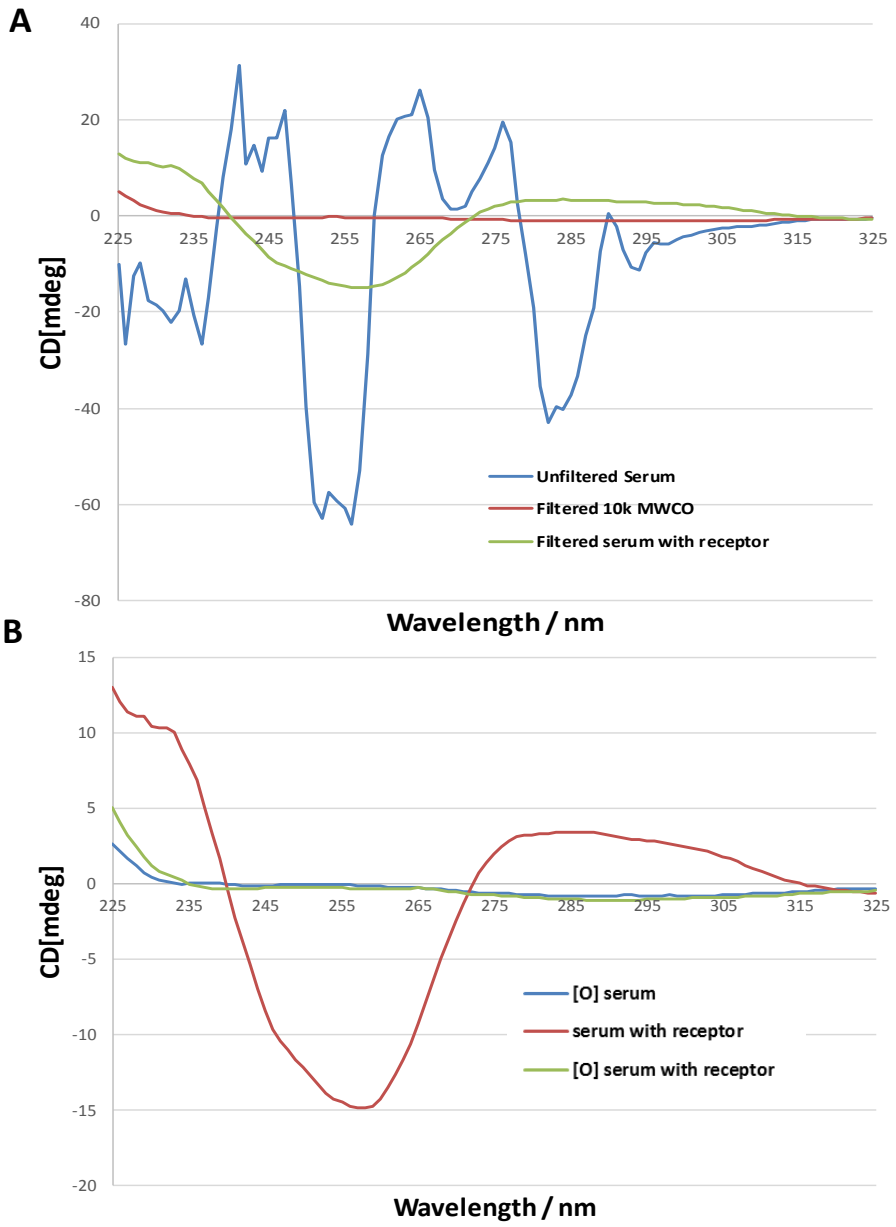
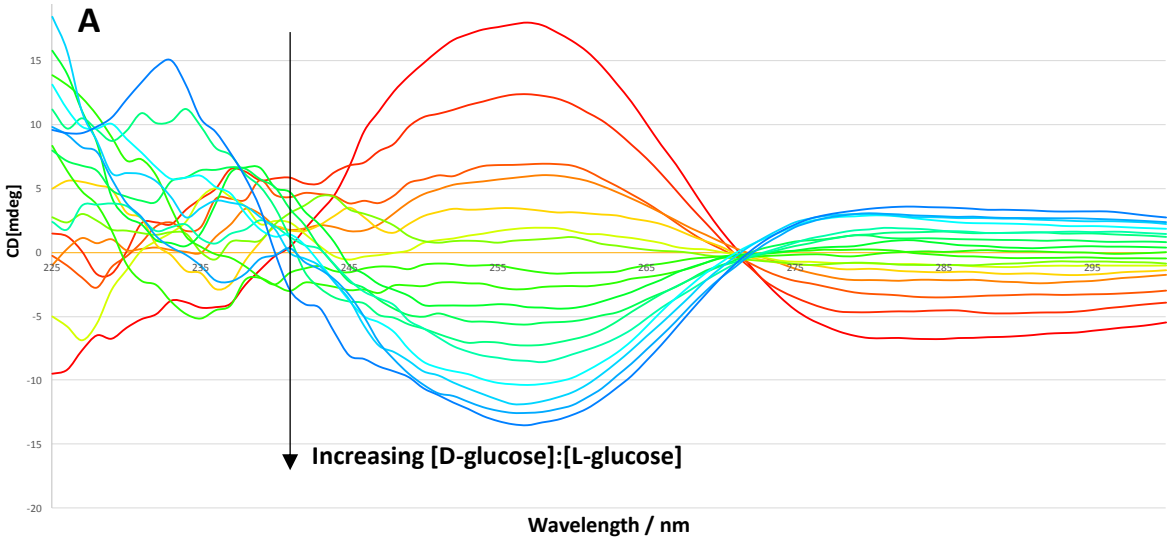
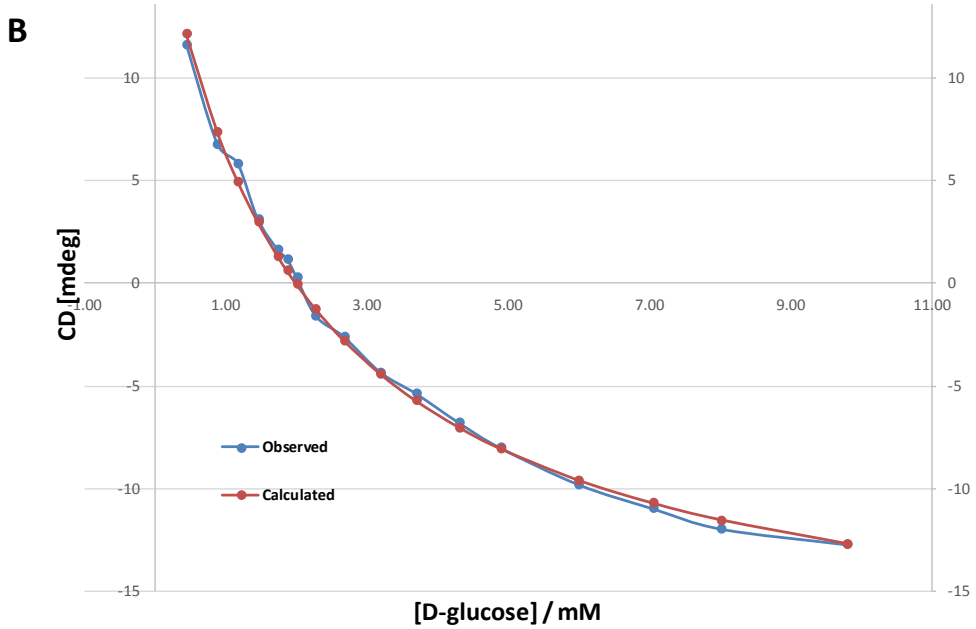


Figure 98 CD absorption spectra from analysis of human serum. The raw unfiltered serum produced very large absorption artefacts, presumably due to high molecular weight components (proteins, antibodies etc.) or light scattering from insoluble particles. Filtration through a membrane (10,000 MWCO) and analysis of this medium gave no observed CD absorption (A). Addition of receptor **3** to this medium did induce a CD signal, due to the binding of **3** to D-glucose. Removal of D-glucose from the human serum (using glucose oxidase and catalase) showed no CD signal upon addition of receptor however (B), which would imply that all D-glucose has been converted to non-binding D-gluconic acid **174**.

With the knowledge that receptor **3** still provides a measurable output upon binding in human serum using CD, attempts were undertaken to ascertain whether this method could be used to quantify the concentration of D-glucose in human serum – a proof of principle of receptor **3** functioning as a D-glucose sensor in physiological media. Initially, a medium was prepared that consisted of human serum (still containing D-glucose) and receptor **3**. A solution of L-glucose, with added receptor **3** in human serum, was then titrated into this medium. As seen before for the similar experiment in water, when the CD absorption intensity is equal to zero, the ratio of D and L-glucose complexes will be equal. Therefore, the concentration of L-glucose added at this point will be equal to the concentration of D-glucose in the human serum. Such an experiment was performed successfully, with zero CD signal observed upon reaching an L-glucose concentration of 5.7-5.9 mM – thus the concentration of D-glucose in the human serum must be 5.7-5.9 mM also.

To further exemplify this concept, a system was developed that could measure any concentration of D-glucose in this sample of human serum – replicating a practical procedure of measuring D-glucose in a blood sample. A calibration study was therefore performed, where the intensity of the CD absorption in human serum was measured at many different concentrations of D-glucose. This was achieved by pre-saturation of receptor **3** with L-glucose (at two different concentrations, 2mM and 8 mM) in human serum. Two systems with differing concentrations of L-glucose were tested to ascertain which concentration of L-glucose would give the most accurate results. Titration of D-glucose, with added L-glucose (at either 2mM or 8 mM) and **3**, into this medium gave the predicted decrease in CD absorption intensity as concentration of D-glucose increased (Figure 99).





C

	Observed CD [mdeg]	[D-glucose] / mM
Run 1	-3.50	5.78
Run 2	-3.17	5.58
Run 3	-3.37	5.70
Average	-3.35	5.69
	Final [D-Glucose] / mM	5.69

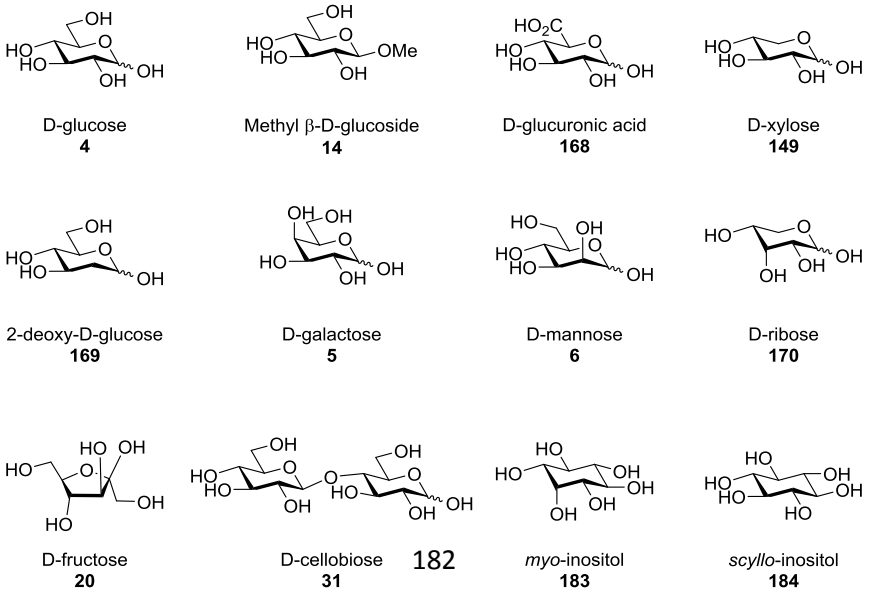
Figure 99 CD results for quantification of D-glucose in human serum. Titration of D-glucose into receptor **3** pre-saturated with L-glucose (here 2 mM), in human serum (D-glucose removed), showed decrease in CD absorption intensity (A). Plotting of CD intensity versus D-glucose concentration gave a calibration curve for D-glucose concentration in this system (B). Measurement of CD signal intensity of receptor **3**, with added L-glucose (2 mM), in human serum (D-glucose still present), allowed quantification of D-glucose in the human serum by inputting this CD signal intensity into the calibration curve equation (C). This gave a measured D-glucose concentration of 5.69 mM in the human serum.

Plotting CD signal intensity against D-glucose concentration for each system (2mM or 8mM L-glucose) gave a curve, and applying a non-linear least squares curve-fitting programme (as for the calculated binding constants) gave the 'calibration curve' and thus equation for the curve. Measurement of CD absorption intensity for receptor **3** in human serum (still containing D-glucose) with added L-glucose (either 2mM or 8mM to replicate each calibration system) was performed 3 times and then averaged. Inputting of this CD intensity to the calibration curve equations afforded D-glucose concentrations of 5.7 mM (2 mM L-glucose system) and 5.8 mM (8 mM L-glucose system) – within the range of measured D-glucose concentration for the human serum (5.7-5.9 mM, discussed earlier). This result would suggest that receptor **3** could be employed to accurately determine D-glucose concentration in physiological systems such as blood – a proof of concept for real world application.

The binding results obtained for **3** and all substrates tested using various techniques (NMR, ITC and CD) are compiled in

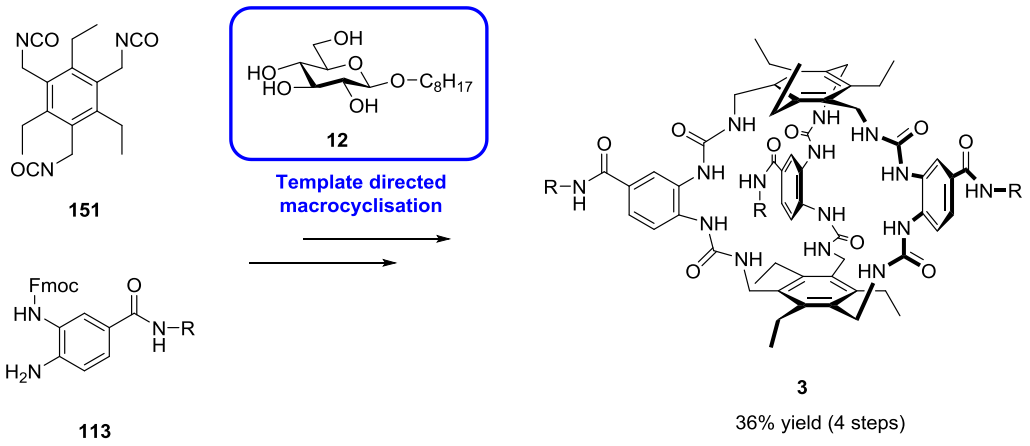
Table **5** below. When listed in this format, the incredible selectivity of **3** for D-glucose **4** is apparent, with most other substrates tested showing no evidence of binding. Those that do exhibit binding affinity do so with affinities 1-3 orders of magnitude weaker, including similar carbohydrates. The only exception is *scyllo*-inositol **184**, essentially the ideal substrate for receptor **3**, and appears to approximately bind an order of magnitude stronger than D-glucose **4**. Fortunately, *scyllo*-inositol **184** is not commonly found in nature, and so selectivity for D-glucose **4** versus biologically relevant substrates remains excellent.

Table 5 Summary of binding results for receptor **3** with various substrates and their associated errors for each technique used (NMR, ITC or CD). All binding studies conducted in H_2O (or D_2O for NMR), at pH 7.4 (with 10 mM phosphate buffer) and at 298 K unless otherwise stated. [a] n.d. = not determinable due to intermediate exchange binding kinetics. [b] No errors given for *scyllo*-inositol due to the inherently large errors obtained during measuring the affinity (described earlier). Substrates that bind to receptor **3** are displayed graphically below for reference.



Substrate (medium)	Binding constant K_a (M^{-1}) determined by:		
	NMR	**ITC**	**CD**
D-Glucose	18,000 ± 1.0%	18,400 ± 2.2%	17,200 ± 9.8%
L-Glucose	-	-	17,200 ± 2.6%
D-Glucose pH 6 (PBS)	-	17,800 ± 3.8%	-
D-Glucose pH 7 (PBS)	-	17,300 ± 5.5%	-
D-Glucose pH 8 (PBS)	-	18,300 ± 1.8%	-
D-Glucose (human serum)	-	11,300 ± 2.4%	10,300 ± 5.9%
D-Glucose (DMEM cell culture)	-	5640 ± 2.1%	-
D-Glucose (DMEM salt control)	-	5160 ± 5.9%	-
D-Glucose (Leibovitz's L-15)	-	5210 ± 8.6%	-
Methyl β-D-Glucoside	7520 ± 5.5%	7890 ± 16%	-
D-Glucuronic Acid	n.d.[a]	5340 ± 3.5%	4400 ± 4.3%
D-Xylose	n.d.[a]	5800 ± 3.0%	5000 ± 8.8%
2-Deoxy-D-Glucose	n.d.[a]	725 ± 5.7%	680 ± 5.7%
D-Galactose	132 ± 10%	182 ± 2.3%	-
D-Mannose	140 ± 1.3%	143 ± 1.1%	-
D-Ribose	267 ± 3.8%	216 ± 1.9%	-
D-Fructose	51 ± 5.5%	60 ± 2.7%	-
D-Cellobiose	31 ± 9.0%	30 ± 16%	-
Mannitol	-	0	-
D-Gluconic acid	0	0	-
Methyl α-D-Glucoside	0	0	-
N-Acetyl-D-glucosamine	0	0	-
D-Maltose	-	0	-
L-Fucose	-	0	-
Ascorbic Acid	-	0	-
Uracil (PBS)	-	0	-
Uric Acid (PBS)	-	0	-
Cytosine	-	0	-
Adenosine	-	0	-
L-Phenylalanine	-	0	-
L-Tryptophan	-	0	-
Paracetamol	-	0	-
Myo-inositol	7300 ± 7.4%	7600 ± 4.2%	-
Scyllo-inositol[b]	$5\text{-}8 \times 10^5$	3×10^5	-

5.5 Conclusions



Scheme 86 Synthesis of hexaurea macrocycle **3** from components *tris*-isocyanate **151** and linker **113**. The key step was macrocyclisation, where octyl β-D-glucoside **12** appeared to template the formation of the target macrocycle in ~50% yield. Global deprotection of the solubilising groups and neutralisation to pH 7.4 yielded the water soluble macrocycle **3**. R = G2MM solubilising group.

The potential of the urea spacer unit was fully realised when incorporated into triethylbenzene hexaurea **3**. Synthesis of **3** was achieved from components **151** and **113** in 4 steps, in 36% yield (Scheme 86). The key synthetic step was the macrocyclisation, which was templated in the presence of octyl β-D-glucoside **12**, providing the protected macrocycle in ~50% yield. Deprotection of the G2MM solubilising groups afforded the water soluble macrocycle **3**.

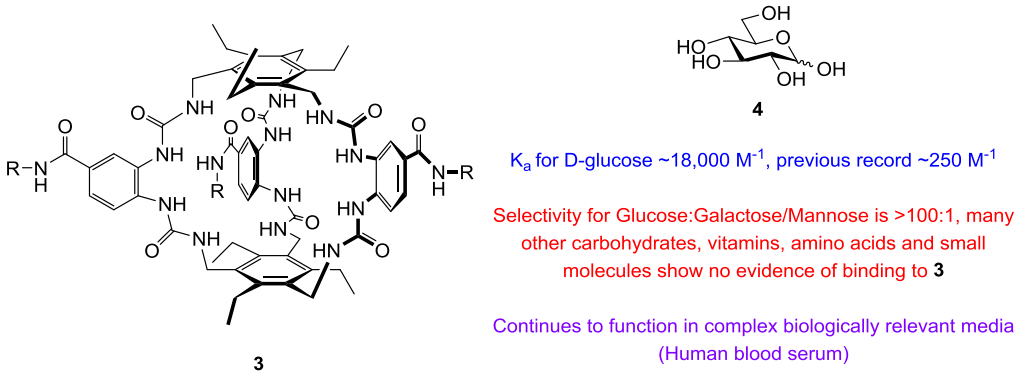


Figure 100 Hexaurea **3** shows unprecedented binding affinity for D-glucose **4**, and extraordinary selectivity for D-glucose **4** against other small biologically relevant molecules (including other similar carbohydrates). Receptor **3** even continues to selectively recognise D-glucose **4** in human blood serum, which allowed quantification of D-glucose **4** concentrations in the human serum using CD – a proof of concept of **3** functioning as a D-glucose **4** sensor.

Hexaurea **3** showed unprecedented affinity for D-glucose **4**, with an association constant of K_a ~18,000 M^{-1} – the previous record for a synthetic lectin system based on non-covalent interactions being K_a ~250 M^{-1} (Figure 100). More impressive however, is the extraordinary selectivity that receptor **3** demonstrates for D-glucose **4** over many other biologically relevant molecules. These include very similar carbohydrates D-galactose **5** and D-mannose **6**, which are bound >100 times weaker. Many other substrates tested, including other carbohydrates, vitamins, amino acids, medically important substrates (such as paracetamol **181** and vitamin C **182**), small aromatics and other biologically relevant molecules showed no evidence of binding to **3** when tested. This exceptional selectivity of **3** for D-glucose **4** allowed it to continue to function as a receptor in biologically relevant media – such as human serum derived from whole blood. This allowed quantification of D-glucose **4** in the human serum using optical methods (CD), a proof of concept of receptor **3** functioning as a D-glucose sensor for potentially medicinal purposes.

The affinity and selectivity receptor **3** demonstrates for D-glucose **4** are unparalleled by any current synthetic system, and are more closely related to biomolecules such as proteins – even then receptor **3** outperforms many naturally occurring carbohydrate binding molecules found in nature by affinity but most importantly by also demonstrating vastly superior selectivity. Based on this, receptor **3** can be confidently described as 'biomimetic' for its target substrate: D-glucose **4**. In terms of practical application, receptor **3** is easy to synthesise, stable and shows no toxicity when tested. Given its synthetic nature, it should be readily adaptable and derivatisable for purposes towards treatments for Diabetes: continuous glucose monitoring and glucose-responsive insulins.

Future work surrounding receptor **3** is predicted to be very rich, with many derivatives of the structure possible due to the modular synthetic approach. Variations of the surfaces and urea spacer units would lead to a library of receptors with the potential to tune the affinities and selectivities for certain substrates accordingly. It could be possible that other substrates could be targeted through these synthetic modifications, for example variation of one of the urea spacers may increase cavity size and allow targeting of larger substrates such a *N*-acetyl-D-glucosamine **7**. Incorporation of fluorescent components may allow a fluorescence output upon substrate binding, increasing potential of receptors like **3** to be applied towards medicinal sensing applications (i.e. continuous glucose monitoring). Given that the protected receptor **3** is soluble in apolar media, it could also be applied to the replication of glucose transporter proteins, by shuttling D-glucose **4** across membranes – something largely unprecedented by a synthetic receptor.

Chapter 6 – Conclusions

The main body of work discussed within this thesis is the design, synthesis and analysis of carbohydrate receptors that incorporate ureas into the spacer units. This approach is distinct from previous receptor designs which employed a bis-amide spacer unit. By progressing to a bis-urea spacer, it was predicted that improved selectivities and affinities would be achieved for target carbohydrate guests – through increased contributions from hydrogen bonding towards affinities and increased cavity sizes disfavouring binding to off-target substrates (such as aromatic substrates).

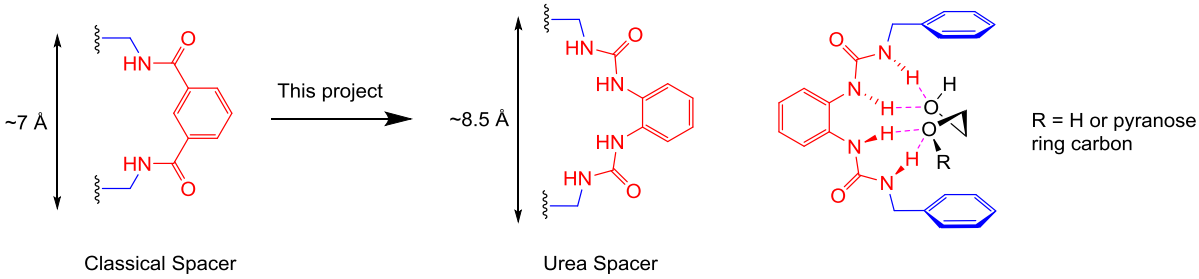


Figure 101 Evolution of bis-amide spacer unit to bis-urea spacer unit (left). The bis-urea features double the number of hydrogen bond donors as well as increased distance between the aromatic surfaces. Hypothesised hydrogen bonding array provided by the bis-urea spacer towards a diol motif commonly found in carbohydrates (right).

The first receptor design to be pursued was anthracene tetraurea **36** (Chapter 2), which was designed to target linear polysaccharides derived from D-cellulose **8** and D-chitin **9**. The anthracene surfaces were designed to run parallel to polysaccharide chain to provide CH-π interactions with the axial CHs from the sugar, but also to provide a fluorescence response upon binding. The bis-urea spacer was also predicted to form hydrogen bonding interactions with the hydroxyls of the sugar by MCMM modelling studies. The design of anthracene tetraurea **36** features water solubilising groups attached to the anthracene. This was to facilitate development of synthetic methodology of the bis-urea linkage within the macrocycle. Synthesis of protected macrocycle **37** was achieved in two steps from phenylene 1,2-diamine and anthracene isocyanate **39**. The alkoxy carboxyl groups on the anthracenes proved to be a great hindrance however, as they made the anthracene surfaces very electron rich and thus very susceptible to irreversible oxidation. The electron rich anthracenes also facilitated fragmentation of the macrocycle upon acidic deprotection of the solubilising groups, meaning no route to the water soluble macrocycle **36** was possible.

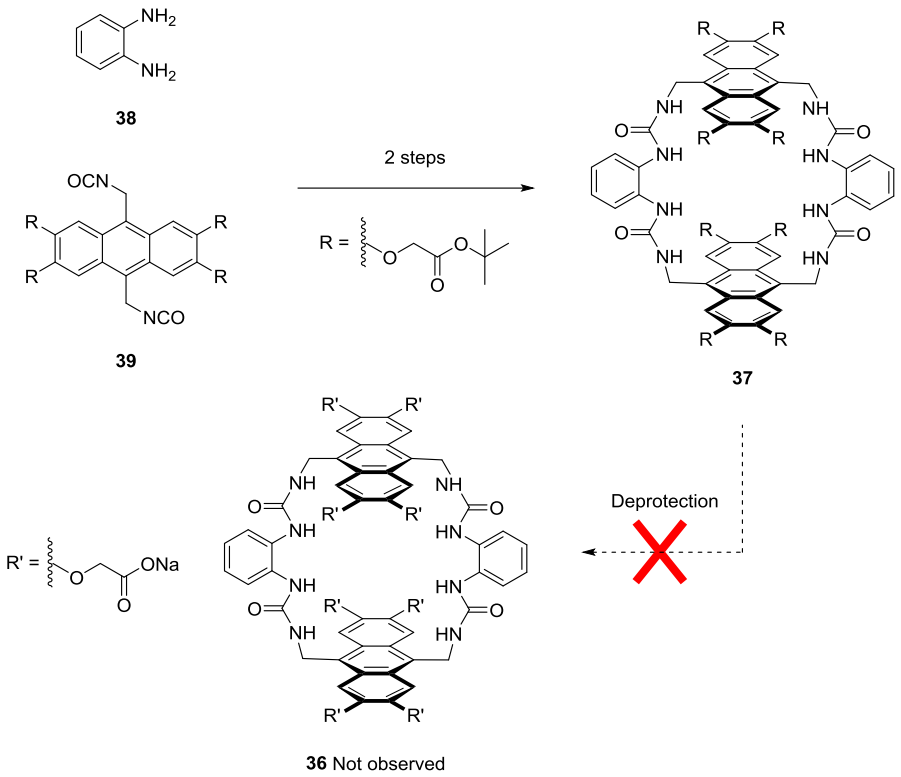


Figure 102 Successful methodology (Scheme 12 and Scheme 14) developed to access tetraurea macrocycle **37** from diamine **38** and bis-isocyanate **39**. Deprotection to water soluble macrocycle **36** was not achieved.

An alternate anthracene tetraurea design was then proposed: one which featured a much simpler anthracene core and solubilising groups attached to the urea spacer units. It was predicted that this approach would circumvent the issues observed for previous macrocycle **37**. This was preceded by successful development of linker unit **113** (Chapter 3), which featured a large solubilising group (G2MM). Incorporation of this linker unit into an anthracene macrocycle afforded protected macrocycle **120** (Chapter 4). However, tetraurea **120** still suffered from similar stability problems as for **37** but ultimately proved much more controllable. This allowed controlled deprotection of the solubilising groups and access to water soluble anthracene receptor **1**. Receptor **1** showed good affinities and selectivities for its target substrates of linear cellodextrins, such as D-cellotriose **2**, in binding studies by NMR , ITC and fluorescence. The affinities demonstrated by **1** are comparable to natural lectins but do not match previous pyrene receptor **32**. Nevertheless, tetraurea **1** is much more synthetically accessible than **32** and could improve its binding affinity through synthetic modification, with the modular synthesis enabling several points for derivatisation. Receptor **1** also demonstrated a fluorescence response upon binding, and this could lead to application of **1** in the detection of β-glucans. Work is currently ongoing to fully characterise the binding affinities and selectivity of receptor **1**.

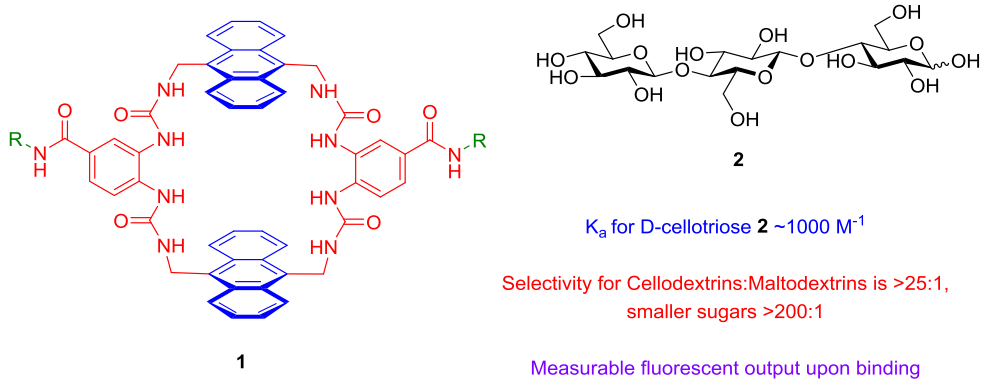
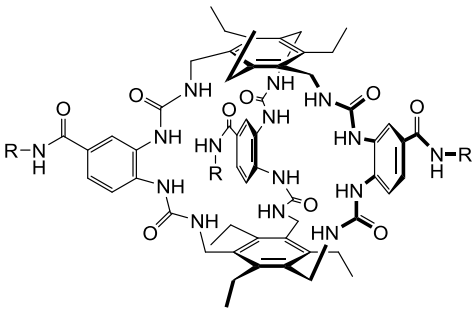
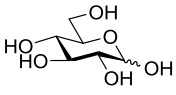


Figure 103 Anthracene tetraurea **1** was successfully synthesised and showed a distinct preference for the binding of linear cellodextrins, such as D-cellotriose **2**, which were bound selectively over monosaccharides and non-linear oligosaccharides (i.e. maltodextrins). Receptor **1** also produced a measurable fluorescent output upon binding, which could enable **1** to be employed as a sensor for the detection of β-glucans.

The potential of the urea spacer unit was fully realised when incorporated into triethylbenzene hexaurea **3**, which featured none of the instability issues possessed by tetraurea anthracene macrocycles **37** and **1**, as well as being much easier to synthesise. Hexaurea **3** showed unprecedented affinity for D-glucose **4**, with a >70 times increased affinity over the previous record, which is attributed to urea spacer providing a much denser and more complementary arrangement of hydrogen bond donors. More impressive however, is the extraordinary selectivity that receptor **3** demonstrates for D-glucose **4** over many other biologically relevant molecules. These include similar carbohydrates that are bound ~100 times weaker. Most other substrates tested, including other carbohydrates, vitamins, amino acids, small aromatics, drugs and other biologically relevant compounds showed no evidence of binding to receptor **3**. This exceptional selectivity for D-glucose **4** allowed **3** to continue to function in complex biological media (such as human blood serum), allowing accurate quantification of D-glucose **4** in the human serum using optical methods – a proof of concept of **3** functioning as a D-glucose **4** sensor. The performance of hexaurea **3** far exceeds any previous synthetic system and even outcompetes most carbohydrate binding proteins (such as lectins), therefore making hexaurea **3** the first synthetic biomimetic receptor for D-glucose **4** based on non-covalent interactions and thus a true synthetic lectin.



3



4

K_a for D-glucose ~18,000 M^{-1}, previous record ~250 M^{-1}

Unprecedented selectivity for D-glucose: many other carbohydrates, vitamins, amino acids and small molecules show no evidence of binding to **3**

Continues to function in complex biologically relevant media (Human blood serum). **3** shown to be able to accurately quantify D-glucose concentration in human serum with optical methods.

Figure 104 Hexaurea **3** shows unprecedented binding affinity for D-glucose **4**, and extraordinary selectivity for D-glucose **4** against other small biologically relevant molecules (including other similar carbohydrates). Receptor **3** even continues to selectively recognise D-glucose **4** in human blood serum, which allowed quantification of D-glucose **4** concentrations in the human serum using optical methods

Chapter 7 – Experimental

7.1 General Experimental, materials and methods

Commercial reagents were purchased from Sigma-Aldrich, Alfa-Aesar, Fisher Scientific, VWR, Frontier Scientific, Fluorochem, Carbosynth, Megazyme, Santa Cruz Biotech, TCI chemicals or Acros Organics and were used without further purification unless otherwise specified. Cell culture media were obtained from ThermoFisher Scientific. Human Serum and enzymes were obtained from Sigma Aldrich. All air and moisture sensitive manipulations were carried out using standard vacuum line and Schlenk techniques. Solvents for air and moisture sensitive manipulations were obtained from an Anhydrous Engineering Solvent Purification System, distilled and dried over activated molecular sieves, or purchased from Acros Organics.

Flash column chromatography was performed using silica gel 60 Å (Sigma Aldrich, particle size 35-70 micron) and a suitable eluent. Reverse phase flash chromatography was performed on a Biotage Isolera One with Biotage SNAP Ultra C18 25 µm columns, typically with acetone/water eluents. TLC was performed using aluminium backed TLC plates (Merck-Keiselgel 60 F_{254}) and visualised using UV fluorescence (254 or 365 nm) and/or developed using ninhydrin, potassium permanganate, EtOH/H_2SO_4, vanillin, $Pd(OAc)_2/H_2O$ or iodine.

HPLC chromatography was performed using a Waters 600 Controller with a Waters 2998 Photodiode Array Detector. For analytical runs a XSELECT CSH C18 5 µm (4.6 × 150 mm) column was used and for preparative runs a XSELECT CSH Prep C18 5 µm OBD (19 × 250 mm) column was used, normally with an acetone/water or methanol/0.1% trifluoracetic acid (aqueous) solvent mixtures.

^{1}H and ^{13}C NMR spectra were recorded on Varian VNMR 400 MHz, Bruker 400 MHz, Varian VNMRS 500 MHz, Bruker Advance III HD Cryo 500 MHz and Varian VNMRS (equipped with a ^{1}H observe cryogenically cooled probe) 600 MHz spectrometers. All spectra were obtained at ambient temperature unless stated otherwise. All ^{1}H and ^{13}C NMR chemical shifts are reported relative to the ^{1}H (residual) and ^{13}C chemical shifts of the solvent as standard.

IR spectra were recorded on a Perkin-Elmer Spectrum One FT-IR spectrometer with an ATR accessory and frequencies reported in wavenumbers (cm^{-1}). ESI-HRMS (electrospray ionisation high resolution mass spectrometry) was performed on a Bruker Daltonics micrOTOF II.

UV/vis spectra were recorded on a Thermo Helios Gamma or an Agilent Cary 300 spectrometer. Fluorescence spectra were recorded on a PerkinElmer LS45. All optical measurements were conducted in Ultrapure water using a Hellma QS-101 or QS-111 four-sided quartz cuvette. Any in-cuvette stirring was conducted using Fisherbrand PTFE encased 5mm x 2mm micro magnetic stir bars. Isothermal

Titration Calorimetry (ITC) titrations were conducted using a Malvern Microcal iTC200. Circular Dichroism was performed on a Jasco J-815 Circular Dichroism Spectropolarimeter using Hellma QS-101 quartz absorption cells (1 mm path length).

7.2 General procedures for binding studies

^{1}H NMR titrations. 1H-NMR titrations were performed on a Varian VNMRS 600 MHz spectrometer equipped with a ^{1}H observe cryogenically cooled probe. Solutions of saccharides in D_2O (99.9%) with 10 mM phosphate buffer (pH 7.4), containing receptor at a known concentration to be used in the experiment, were prepared and allowed to equilibrate overnight before use if necessary. Aliquots were then added to an NMR tube containing a known concentration of receptor solution (typically 50 μM – 250 μM). The receptor concentration was therefore held constant while the carbohydrate concentration was increased. The NMR tube was shaken after each addition, manually centrifuged and the ^{1}H-NMR spectra acquired at 298 K.

Isothermal Titration Microcalorimetry (ITC) titrations. Isothermal Titration MicroCalorimetry (ITC) experiments were performed on a MicroCal iTC200 microcalorimeter. ITC experiments were carried out at 298 K. Saccharide solutions were prepared in HPLC-grade water with 10 mM phosphate buffer solution (pH 7.4) and allowed to equilibrate overnight if necessary. The sample cell was charged with a known concentration of receptor solution in HPLC-grade water with 10 mM phosphate buffer solution at pH 7.4 (typically 50 μM – 200 μM). Then, aliquots (typically 1.0 μL) of carbohydrate solution were added and the evolution of heat was followed as a function of time. Heats of dilution were measured by injecting the same carbohydrate solution into HPLC-grade water with 10 mM phosphate buffer solution at pH 7.4, using identical conditions. For every addition, the heat of dilution was subtracted from the heat of binding using a MicroCal software programme implemented in ORIGIN 7.0. This gave an XY matrix of heat vs. total guest concentration. This matrix was then imported into a specially written Excel programme to fit the data to a 1:1 binding model to give a K_a.

Preparation of biological media. Cell culture media were obtained from ThermoFisher Scientific and human serum from Sigma-Aldrich, and were aliquoted into 10 mL batches in a specially designed sterile fume hood. The cell culture media and human serum were passed through a membrane (10k MWCO) and then with 10 mM phosphate buffer (pH 7.4). These batches were then used immediately for ITC binding studies to avoid contamination. The two varieties of cell culture media used were: Dulbecco's Modified Eagle Medium (DMEM, no glucose, cat. no. 11966025) and Leibovitz's L-15 Medium (no glucose, cat. no. 11415049). The Human Serum used originates from male AB clotted whole blood (cat. no. H6914). Formulations for both cell culture media, phosphate buffered saline (PBS) and the DMEM salt control media are listed in the tables below

Removal of D-glucose from human serum. The concentration of initial D-glucose present in the human serum was measured to be 6.2 mM using a YSI 2300 STAT Plus Glucose and L-Lactate analyser. The D-glucose was oxidised to the D-glucono-δ-lactone (which then hydrolysed to gluconic acid in solution at pH 7.4) using a combination of enzymes: glucose oxidase and catalase. After oxidation, the concentration of D-glucose in the human serum was measured to be 0-0.005 mM (using the YSI 2300 STAT Plus). The procedure for the oxidation of D-glucose in the serum is as follows: A solution of enzymes (20 mL) consisting of glucose oxidase (10,000 U) and catalase (300,000 U) was prepared and then dialysed (500 MWCO) to remove low MW contaminants (notably glycerol). 0.2 mL of this enzyme solution was added to human serum (10 mL) at 25-30 °C. The mixture was stirred at 25-30 °C while air was bubbled through the suspension for 2 hours. The mixture was then cooled to room temperature and passed through a membrane (10k MWCO) to give a colourless solution, which was used immediately for the corresponding binding studies.

Table 6 Formulation for phosphate buffered saline (PBS).

Component	**Concentration (mM)**
Sodium Chloride (NaCl)	140
Potassium Chloride (KCl)	2.7
Disodium phosphate (Na_2HPO_4)	10
Monopotassium phosphate (KH_2PO_4)	1.8

Table 7 Formulation for DMEM salt control

Component	**Concentration (mM)**
Calcium Chloride ($CaCl_2$) (anhyd.)	1.8
Ferric Nitrate ($Fe(NO_3)_3.9H_2O$)	2.50×10^{-4}
Magnesium Sulfate ($MgSO_4$) (anhyd.)	0.81
Potassium Chloride (KCl)	5.3
Sodium Bicarbonate ($NaHCO_3$)	44
Sodium Chloride (NaCl)	110
Sodium Phosphate monobasic (NaH_2PO_4-H_2O)	0.9

Table 8 Formulation for DMEM cell culture medium used, as described by the manufacturer.

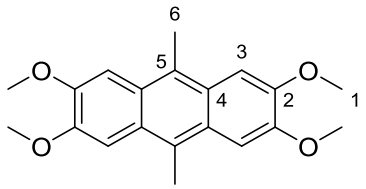
Component	**Concentration (mM)**
Glycine	0.4
L-Arginine hydrochloride	0.4
L-Cystine 2HCl	0.2
L-Glutamine	4.0
L-Histidine hydrochloride-H_2O	0.2
L-Isoleucine	0.8
L-Leucine	0.8
L-Lysine hydrochloride	0.8
L-Methionine	0.2
L-Phenylalanine	0.4
L-Serine	0.4
L-Threonine	0.8
L-Tryptophan	0.078
L-Tyrosine disodium salt dihydrate	0.4
L-Valine	0.8
Choline chloride	0.029
D-Calcium pantothenate	0.0084
Folic Acid	0.0091
Niacinamide	0.033
Pyridoxine hydrochloride	0.019
Riboflavin	0.0011
Thiamine hydrochloride	0.012
i-Inositol	0.04
Calcium Chloride ($CaCl_2$) (anhyd.)	1.8
Ferric Nitrate ($Fe(NO_3)_3.9H_2O$)	2.50×10^{-4}
Magnesium Sulfate ($MgSO_4$) (anhyd.)	0.81
Potassium Chloride (KCl)	5.3
Sodium Bicarbonate ($NaHCO_3$)	44
Sodium Chloride (NaCl)	110
Sodium Phosphate monobasic (NaH_2PO_4-H_2O)	0.91
Phenol Red	0.040

Table 9 Formulation for Leibovitz's L-15 cell culture medium used, as described by the manufacturer.

Component	**Concentration (mM)**
Glycine	2.7
L-Alanine	2.5
L-Arginine	2.9
L-Asparagine	1.9
L-Cysteine	1.0
L-Glutamine	2.1
L-Histidine	1.6
L-Isoleucine	1.9
L-Leucine	0.95
L-Lysine	0.51
L-Methionine	0.50
L-Phenylalanine	0.76
L-Serine	1.9
L-Threonine	2.5
L-Tryptophan	0.098
L-Tyrosine	1.6
L-Valine	0.85
Choline chloride	0.0071
D-Calcium pantothenate	0.0021
Folic Acid	0.0023
Niacinamide	0.0082
Pyridoxine hydrochloride	0.0049
Riboflavin 5'-phosphate Na	2.0×10^{-4}
Thiamine monophosphate	0.0022
i-Inositol	0.011
Calcium Chloride ($CaCl_2$) (anhyd.)	1.3
Magnesium Chloride (anhydrous)	0.99
Magnesium Sulfate ($MgSO_4$) (anhyd.)	0.81
Potassium Chloride (KCl)	5.3
Potassium Phosphate monobasic (KH_2PO_4)	0.44
Sodium Chloride (NaCl)	140
Sodium Phosphate dibasic (Na_2HPO_4) anhydrous	1.3
D-Galactose	5.0
Phenol Red	0.027
Sodium Pyruvate	5.0

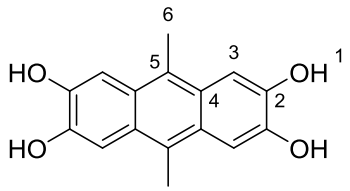
7.3 Synthetic procedures and characterisation

2,3,6,7-tetramethoxy-9,10-dimethylanthracene (46)[118]



A solution of concentrated H_2SO_4 (24 mL, 95% aqueous) was cooled to -10 °C. A mixture of 1,2-dimethoxybenzene **47** (5.4 mL, 42.4 mmol), acetaldehyde **48** (2.4 mL, 42.4 mmol) and acetonitrile (6 mL) was added *via* syringe pump over 2 hours, ensuring the temperature did not exceed 5 °C. The reaction mixture was then stirred under N_2 at -10 °C for 30 mins and then warmed to room temperature, after which the mixture was stirred for 1 hour. The reaction mixture was then poured onto ice and water (160 mL) added. The precipitate was then filtered, washed with water and air-dried. The crude solid was then purified by flash column chromatography (100% CH_2Cl_2) to yield **46** (3.8 g, 11.6 mmol, 55%) as an off white solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 2.95 (s, 6H, 2 × C(6)H$_3$), 4.08 (s, 12H, 4 × OC(1)H$_3$), 7.41 (s, 4H, 4 × C(3)H), [lit.[118] (400 MHz, $CDCl_3$) δ 2.95, 4.08, 7.40]; **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 15.1 (C(6)H_3), 56.0 (OC(1)H_3), 102.6 (C(3)H), 124.2 (C(5)), 126.0 (C(6)), 148.9 (C(6)). **ν_{max}** 3019, 2996, 2963, 2833, 1495, 1164, cm^{-1}; **HRMS:** (ESI$^+$) Calculated for $C_{20}H_{22}O_4Na$: 349.1410, found $[M+Na]^+$: 349.1405.

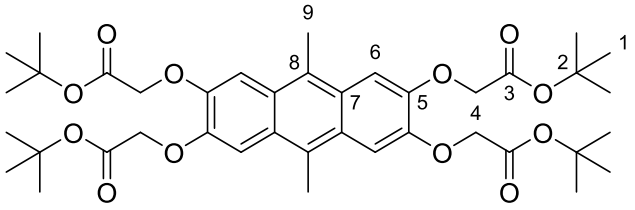
9,10-dimethylanthracene-2,3,6,7-tetraol (45)[119]



Under an inert N_2 atmosphere, **46** (3.8 g, 11.6 mmol) was suspended in anhydrous CH_2Cl_2 (600 mL) and BBr_3 (6.4 mL, 1M in CH_2Cl_2) was added dropwise over 5 mins. The reaction mixture was then stirred under reflux for 16 hours, after which it was cooled to room temperature and poured into a solution of HCl (400 mL, 1M) cooled to 0 °C and stirred for 30 mins. The precipitate was then extracted with EtOAc (4 × 350 mL), washed with water (200 mL, brine (200 mL) and dried ($MgSO_4$). The solvent was removed under vacuum and the crude solid purified by flash column chromatography (5% MeOH:CH_2Cl_2 → 10% MeOH:CH_2Cl_2) to yield **45** (2.5 g, 9.16 mmol, 79%) as a deep red solid. **^{1}H NMR:** (400 MHz, ($(CD_3)_2CO$): δ 2.78 (s, 6H, 2 × C(6)H$_3$), 7.50 (s, 4H, 4 × C(3)H), 8.34 (s, 4H, OH), [lit.[119] (400 MHz, $(CD_3)_2SO$) δ 2.71, 7.40, 9.45]; **^{13}C NMR:** (100 MHz, ($(CD_3)_2CO$): δ 13.6 (C(6)H_3), 105.5 (C(3)H), 122.2

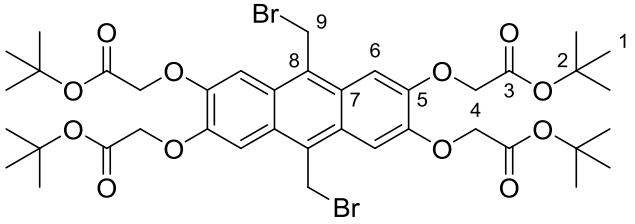
(C(5)), 126.1 (C(4)), 145.4 (C(2)). **ν_{max}** 3210, 3020, 2961, 2837, 1501, 1168 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{16}H_{14}O_4Na$: 293.0784, found $[M+Na]^+$: 293.0792.

tetra-*tert*-butyl-2,2',2'',2'''-((9,10-dimethylanthracene-2,3,6,7-tetrayl)tetrakis(oxy))tetraacetate (44)



Under an inert N_2 atmosphere, **45** (2.35 g, 8.7 mmol) was dissolved in anhydrous THF (500 mL). K_2CO_3 (4.9 g, 35.2 mmol) and *tert*-butyl bromoacetate (7 mL, 47.4 mmol) were added and the reaction mixture stirred under reflux for 16 hours. The mixture was cooled to room temperature and the solvent removed under vacuum. The crude residue was then dissolved in CH_2Cl_2 (500mL) and washed with water (150 mL), brine (200 mL) and dried ($MgSO_4$). The solvent was removed under vacuum and the crude residue purified by flash column chromatography (1% MeOH:CH_2Cl_2) to yield **46** (3.8 g, 5.2 mmol, 60%) as a yellow solid**. ^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 1.49 (s, 36H, 3 × C(1)H$_3$), 2.85 (s, 6H, 2 × C(9)H$_3$), 4.76 (s, 8H, C(4)H$_2$), 7.38 (s, 4H, 4 × C(6)H); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 14.6 (C(9)H_3), 28.1 (C(1)H_3), 66.6 (C(4)H_2), 82.3 (C(2)$(CH_3)_3$), 105.7 (C(6)H), 124.3 (C8), 126.2 (C7), 147.2 (C5), 167.8 (C(3)O); **ν_{max}** 2987, 2901, 1750, 1453, 1369, 1145, 1066 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{40}H_{54}O_{12}Na$: 749.3507, found $[M+Na]^+$: 749.3520.

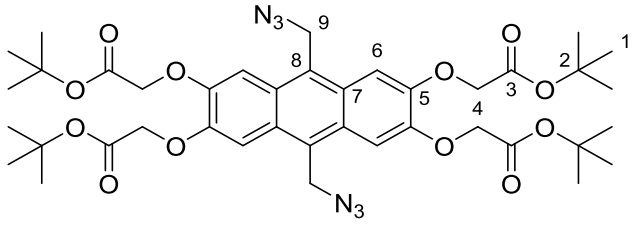
tetra-*tert*-butyl-2,2',2'',2'''-((9,10-bis(bromomethyl)anthracene-2,3,6,7-tetrayl)tetrakis(oxy)) tetraacetate (43)



Under an inert N_2 atmosphere, **44** (3 g, 4.1 mmol) was dissolved in anhydrous CH_2Cl_2 (500 mL). NBS (1.84 g, 10.3 mmol) and ABCN (50 mg, 5 mol%) were added, and the mixture stirred under reflux for 1.5 hours. The reaction mixture was then cooled to room temperature and diluted with CH_2Cl_2 (300 mL). The solution was washed with NaOH (300 mL, 1M), water (300 mL) and the solvent removed under vacuum to yield **43** (3.5 g, 4.0 mmol, 98%) as an orange solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 1.52 (s, 36H, 3 × C(1)H$_3$), 4.81 (s, 8H, C(4)H$_2$), 5.22 (s, 4H, 2 × C(9)H$_2$), 7.38 (s, 4H, 4 × C(6)H); **^{13}C NMR:** (100

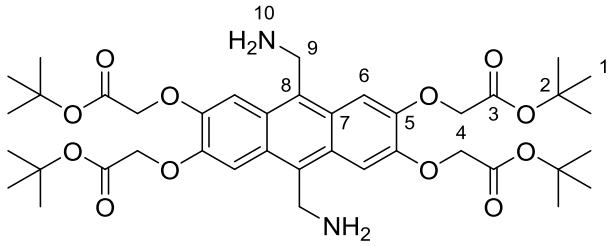
MHz, ($CDCl_3$): δ 28.1 (C(1)H_3), 29.7 (C(9)H_3), 66.4 (C(4)H_2), 82.6 (C(2)$(CH_3)_3$), 104.2 (C(6)H), 126.2 (C8), 126.4 (C7), 148.7 (C5), 167.4 (C(3)O); ν_{max} 2987, 2933, 1706, 1488, 1362, 1228, 1183, 1066 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{40}H_{52}Br_2O_{12}Na$: 905. 1719, 907.1697, found $[M+Na]^+$: 905.1709, 907.1692.

tetra-*tert*-butyl-2,2',2'',2'''-((9,10-bis(azidomethyl)anthracene-2,3,6,7-tetrayl)tetrakis(oxy)) tetraacetate (42)



Under an inert N_2 atmosphere, **43** (3.5 g, 4.0 mmol) was dissolved in anhydrous MeCN (300 mL). NaN_3 (1 g, 15.9 mmol) was added and the reaction stirred under reflux for 3 hours. The reaction mixture was cooled to room temperature and the solvent removed under vacuum. The crude product was dissolved in CH_2Cl_2 (400 mL), washed with water (3 × 100 mL) and the solvent removed under vacuum to yield **42** (3.2 g, 3.9 mmol, 98%) as an orange solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 1.52 (s, 36H, 3 × C(1)H_3), 4.78 (s, 8H, C(4)H_2), 5.08 (s, 4H, 2 × C(9)H_2), 7.40 (s, 4H, 4 × C(6)H) [lit.[44] ($CDCl_3$, 400 MHz): δ 7.41 (4H, s), 5.08 (4H, s), 4.78 (8H, s), 1.52 (36H, s)]; **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 28.1 (C(1)H_3), 46.9 (C(9)H_3), 66.4 (C(4)H_2), 82.6 (C(2)$(CH_3)_3$), 104.6 (C(6)H), 124.2 (C8), 126.8 (C7), 148.7 (C5), 167.5 (C(3)O); ν_{max} 2988, 2931, 2091, 1736, 1498, 1364, 1227, 1186, 1062cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{40}H_{52}N_6O_{12}Na$: 831.3535, found $[M+Na]^+$: 831.3528.

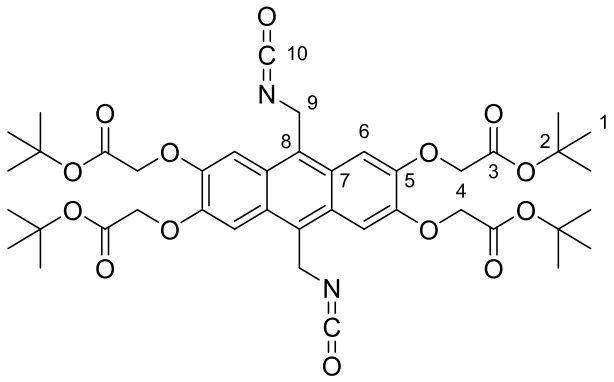
tetra-*tert*-butyl-2,2',2'',2'''-((9,10-bis(aminomethyl)anthracene-2,3,6,7-tetrayl)tetrakis(oxy)) tetraacetate (41



Under an inert N_2 atmosphere, **42** (100 mg, 0.12 mmol) was dissolved in anhydrous degassed THF (8 mL). PMe_3 was added (2.5 mL, 2.5 mmol, 1M in THF) and the mixture stirred at room temperature for 3 hours. Degassed water (2 mL) was added and the reaction mixture stirred for 1 hour. The solvent was then evaporated under a flow of nitrogen and the crude residue dissolved in THF/H_2O (5:1, 3 mL). The solvent was then removed by freeze-drying to yield **41** (90 mg, 0.12 mmol, 96%) as a pale brown solid.

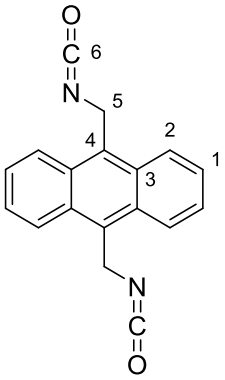
^{1}H NMR: (400 MHz, ($CDCl_3$): δ 1.50 (s, 36H, 3 × C(1)$\underline{H}_3$), 4.58 (s, 4H, 2 × C(9)$\underline{H}_2$), 4.77 (s, 8H, C(4)$\underline{H}_2$), 7.49 (s, 4H, 4 × C(6)$\underline{H}$) [lit.[44] ($CDCl_3$, 400 MHz): δ 7.47 (4H, s),4.73 (8H, s), 4.64 (4H, s), 1.47 (36H, s);]; **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 28.1 ($\underline{C}$(1)H_3), 38.9 ($\underline{C}$(9)H_2), 66.5 ($\underline{C}$(4)H_2), 82.4 ($\underline{C}$(2)$(CH_3)_3$), 105.0 ($\underline{C}$(6)H), 125.8 ($\underline{C}$8), 126.2 ($\underline{C}$7), 148.0 ($\underline{C}$5), 167.6 (C(3)O);); **ν_{max}** 2982, 2926, 1729, 1497, 1358, 1222, 1144, 1069 cm^{-1}; **HRMS**: (MALDI$^+$) Calculated for $C_{40}H_{57}N_2O_{12}$: 757.3906, found [M+H]$^+$: 757.3909.

Tetra-*tert*-butyl-2,2',2'',2'''-((9,10-bis(isocyanatomethyl)anthracene-2,3,6,7-tetrayl)tetrakis(oxy)) tetraacetate (39)



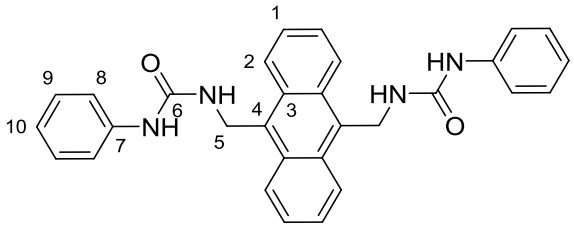
Under an inert N_2 atmosphere, a flask was charged with **41** (30 mg, 0.04 mmol). CH_2Cl_2 (1 mL) and $NaHCO_3$ (aq. 0.1 M, 1 mL) were added, the mixture cooled to 0 °C and rapidly stirred. Triphosgene (9.4 mg, 0.016 mmol) was added and the reaction mixture stirred at room temperature for 30 minutes. The reaction mixture was diluted with CH_2Cl_2 (10 mL) and the organic layer separated, dried ($MgSO_4$) and the solvent removed under vacuum to afford **39** (29 mg, 0.036 mmol, 91%) as an orange solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 1.51 (s, 36H, 3 × C(1)$\underline{H}_3$),), 4.78 (s, 8H, C(4)$\underline{H}_2$), 5.07 (s, 4H, 2 × C(9)$\underline{H}_2$), 7.35 (s, 4H, 4 × C(6)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 28.0 ($\underline{C}$(1)H_3), 39.83 ($\underline{C}$(9)H_2), 66.6 ($\underline{C}$(4)H_2), 82.7 ($\underline{C}$(2)$(CH_3)_3$), 104.4 ($\underline{C}$(6)H), 125.5 ($\underline{C}$8), 126.0 ($\underline{C}$7), 148.7 ($\underline{C}$5), 167.3 ($\underline{C}$(3)O); **ν_{max}** 2979, 2934, 2251, 1734, 1493, 1367, 1225, 1144, 1064 cm^{-1}; **HRMS:** (ESI$^+$) Calculated for $C_{42}H_{52}N_2O_{14}Na$: 831.3311, found [M+Na]$^+$: 831.3319.

9,10-bis(isocyanatomethyl)anthracene (60)



Under an inert N_2 atmosphere, a flask was charged with anthracene-9,10-diyldimethanamine (20 mg, 0.085 mmol). CH_2Cl_2 (1 mL) and $NaHCO_3$ (aq. 0.1M, 1 mL) were added, the mixture cooled to 0 °C and rapidly stirred. Triphosgene (20 mg, 0.068 mmol) was added and the reaction mixture stirred at room temperature for 30 minutes. The reaction mixture was diluted with CH_2Cl_2 (10 mL) and the organic layer separated, dried ($MgSO_4$) and the solvent removed under vacuum to afford **60** (22 mg, 0.076 mmol, 91%) as a yellow solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 5.38 (s, 4H, C(5)$\underline{H}_2$), 7.66 (dd, *J* = 6.9, 3.2 Hz, 4H, C(1)$\underline{H}$), 8.35 (dd, *J* = 6.9, 3.2 Hz, 4H, C(2)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 39.0 ($\underline{C}$5), 124.1 ($\underline{C}$2), 124.2 ($\underline{C}$6), 126.8 (C1), 129.5 ($\underline{C}$3), 129.8 ($\underline{C}$4); **ν_{max}** 2921, 2234, 1491, 1448, 1324, 1185, 858, 751 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{18}H_{12}N_2O_2Na$: 311.0791, found $[M+Na]^+$: 311.0785.

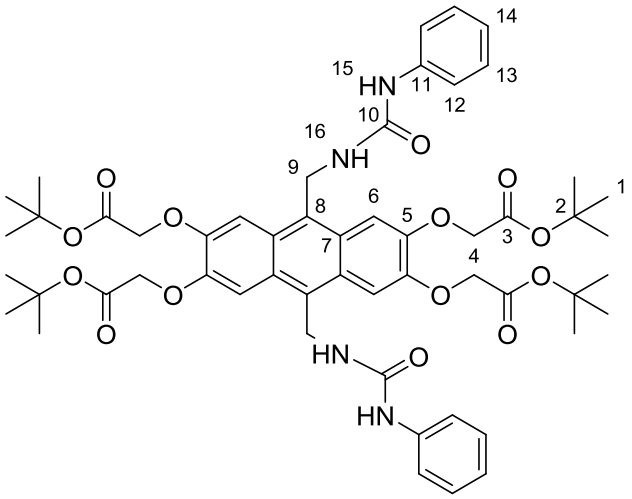
1,1'-(anthracene-9,10-diylbis(methylene))bis(3-phenylurea) (61)[120]



Under an inert N_2 atmosphere, aniline (15 µL, 0.14 mmol) was added to a solution of 60 (20 mg, 0.069 mmol) in dry CH_2Cl_2 (50 mL) and stirred at room temperature for 16 hours. The resultant precipitate was filtered and washed with CH_2Cl_2 to afford **61** (25 mg, 0.053 mmol, 85%) as an off white solid. **^{1}H NMR:** (400 MHz, ($(CD_3)_2SO$): δ 5.31 (d, *J* = 4.8 Hz, 4H, C(5)$\underline{H}_2$), 6.88 (m, 4H, N(12)H and C(10)$\underline{H}$)), 7.21 (d, *J* = 7.4 Hz, 4H, C(8)$\underline{H}$), 6.66-6.70 (m, 2H, C(15)$\underline{H}$), 6.77 (t, *J* = 7.6 Hz, 2H, C(14)$\underline{H}$), 7.38 (d, *J* = 8.9 Hz, 4H, C(2)$\underline{H}$), 7.64 (m, 4H, C(9)$\underline{H}$), 8.53 (m 4H, C(1)$\underline{H}$), 8.62 (s, 2H, N(11)$\underline{H}$),), [lit.[120] (400 MHz, $(CD_3)_2SO$) δ 5.32, 6.88, 7.20, 7.39, 7.63, 8.53, 8.61]; **^{13}C NMR:** (100 MHz, ($(CD_3)_2SO$): δ 36.3 ($\underline{C}$(5)H_2), 118.4 ($\underline{C}$(8)H), 121.8, 125.9, 126.8, 129.4, 130.5 (Aryl-$\underline{C}$), 132.2 ($\underline{C}$3), 141.3 ($\underline{C}$7), 155.8 ($\underline{C}$6), [lit.[120] (100 MHz, $(CD_3)_2SO$)

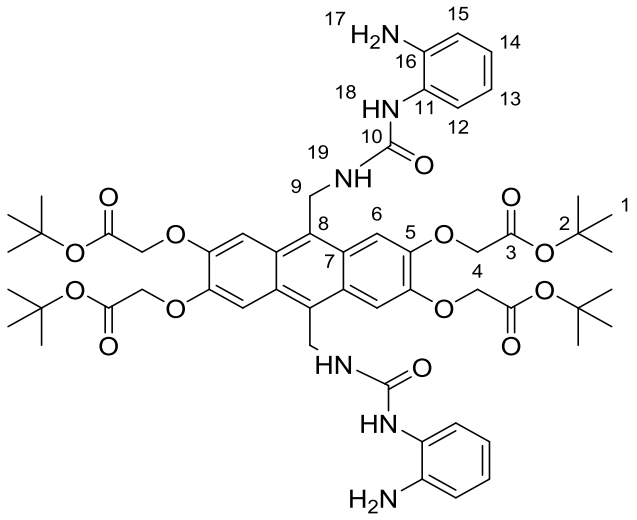
δ 36.2, 118.4, 121.8, 125.9, 126.8, 129.4, 130.5, 132.2, 141.3, 155.8]; **LRMS:** (ESI^+) Calculated for $C_{30}H_{26}N_4O_2Na$: 497.2, found $[M+Na]^+$: 497.2.

***tert*-butyl protected bis-phenylurea anthracene (62)**



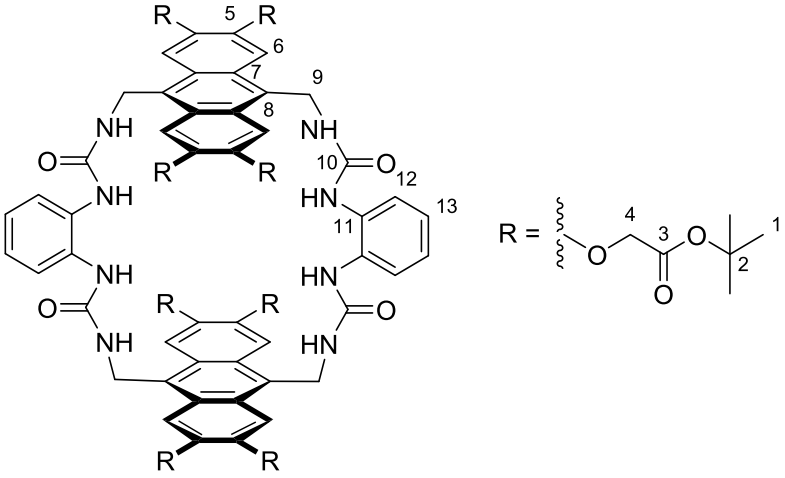
Under an inert N_2 atmosphere, aniline (21 µL, 0.24 mmol) was added to a solution of **39** (20 mg, 0.024 mmol) in dry CH_2Cl_2 (50 mL) and stirred at room temperature for 16 hours. The solvent was removed under vacuum and the crude residue redissolved in EtOAc (30 mL). The organic layer was washed with 1M Acetic acid (20 mL), water (20 mL), brine (20 ml), dried ($MgSO_4$) and the solvent removed under vacuum. The crude solid was purified by flash column chromatography (5% MeOH:CH_2Cl_2 → 10% MeOH:CH_2Cl_2) to afford **62** (20 mg, 0.02 mmol, 82%) as an orange/brown solid. **^{1}H NMR:** (400 MHz, ($(CD_3)_2SO$): δ 1.45 (s, 36H, C(1)H_3), 4.87 (s, 8H, C(4)$\underline{H}_2$), 5.10 (d, *J* = 5.1 Hz, 4H, C(9)$\underline{H}_2$), 6.43 (m, 2H, N(16)$\underline{H}$), 6.88 (t, *J* = 7.3 Hz, 2H, C(14)$\underline{H}$)), 7.20 (t, *J* = 7.8 Hz, 4H, C(13)$\underline{H}$), 7.41 (d, *J* = 8.6 Hz, 4H, C(12)$\underline{H}$), 7.65 (s, 4H, C(6)$\underline{H}$), 8.30 (s, 4H, N(15)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($(CD_3)_2SO$): δ 28.1 (C(1)H_3), 39.7 (C(9)H_2), 66.8 (C(4)H_2), 82.5 (C(2)$(CH_3)_3$), 105.7 (C(6)H), 121.6 (C12), 126.1 (C8), 126.4 (C7), 128.1 (C14), 128.9 (C13), 139.4 (C11), 148.4 (C5), 154.3 (C10), 167.2 (C(3)O); **LRMS:** (ESI^+) Calculated for $C_{54}H_{66}N_4O_{14}Na$: 1017.4, found $[M+Na]^+$: 1017.4.

***tert*-butyl protected half receptor (59)**



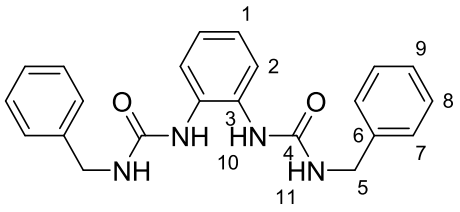
Under an inert N_2 atmosphere, 1,2-phenylene diamine **38** (0.3 g, 2.70 mmol) was dissolved in dry degassed CH_2Cl_2 (120 mL). A solution of **39** (55mg, 0.068 mmol) in dry degassed CH_2Cl_2 (50 mL) was added dropwise over 10 minutes and then stirred at room temperature for 30 minutes. The solvent was removed under vacuum and the crude solid purified by flash column chromatography (80:20 EtOAc:hexane → 5:95 MeOH:CH_2Cl_2 → 10:90 MeOH:CH_2Cl_2) to yield **59** (53 mg, 0.052 mmol, 77%) as an orange brown solid. **^{1}H NMR:** (400 MHz, ($(CD_3)_2SO$): δ 1.47 (s, 36H, 3 × C(1)$\underline{H}_3$),), 4.59 (s, 4H, N(19)$\underline{H}_2$), 4.87 (s, 8H, C(4)$\underline{H}_2$), 5.12 (s, 4H, 2 × C(9)$\underline{H}_2$), 6.50-6.57 (m, 4H, C(13)$\underline{H}$ and N(18)$\underline{H}$), 6.66-6.70 (m, 2H, C(15)$\underline{H}$), 6.77 (t, *J* = 7.6 Hz, 2H, C(14)$\underline{H}$), 7.42-7.48 (m, 4H, C(12)$\underline{H}$ and N(18)$\underline{H}$), 7.66 (s, 4H, 4 × C(6)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($(CD_3)_2SO$): δ 28.2 ($\underline{C}$(1)H_3), 36.7($\underline{C}$(9)H_2), 66.8 ($\underline{C}$(4)H_2), 82.9 ($\underline{C}$(2)$(CH_3)_3$), 105.2 ($\underline{C}$(6)H), 114.5 ($\underline{C}$15), 118.9 ($\underline{C}$13), 122.8 ($\underline{C}$11), 125.2, 125.5 ($\underline{C}$12 and $\underline{C}$14), 125.8 ($\underline{C}$8), 126.2 ($\underline{C}$7), 148.3 ($\underline{C}$5), 149.5 ($\underline{C}$16), 154.2 ($\underline{C}$10), 167.5 ($\underline{C}$(3)O); $\boldsymbol{\nu_{max}}$ 3315, 2973, 2901, 1733, 1622, 1494, 1393, 1225, 1146, 1057, 742 cm^{-1}; **HRMS:** (ESI$^+$) Calculated for $C_{54}H_{68}N_6O_{14}Na$: 1047.4686, found $[M+Na]^+$: 1047.4689.

***tert*-butyl protected tetra-urea macrocycle (37)**



Under an inert N_2 atmosphere, **39** (40 mg, 0.049 mmol) was dissolved in dry degassed CH_2Cl_2 (600 mL). To this was added a solution of **59** (50 mg, 0.049 mmol) in dry degassed pyridine (60 mL) dropwise over 20 minutes. The reaction was stirred at room temperature for 16 hours and then the solvent removed under reduced pressure. The crude residue was suspended in HPLC grade water and freeze-dryed to afford a fine crude solid. The product was then purified by reverse phase HPLC (acetone/water) and freeze dried to afford **37** (50 mg, 0.027 mmol, 56%) as an off white solid. **^{1}H NMR:** (400 MHz, $(CD_3)_2CO$): δ 1.49 (s, 72H, 3 × C(1)$\underline{H}_3$),), 4.78 (m, 16H, C(4)$\underline{H}_2$), 5.12 (m, 8H, C(9)$\underline{H}_2$), 6.95 (s, 4H, C(13)$\underline{H}$), 7.65 (s, 8H, C(6)$\underline{H}$), 8.07 (s, 4H, C(12)$\underline{H}$); **^{13}C NMR:** (100 MHz, $(CD_3)_2CO$): δ 27.3 (C(1)H_3), 36.0 (C(9)H_2), 66.0 (C(4)H_2), 81.6 (C(2)$(CH_3)_3$), 105.3 (C(6)H), 126.3 (C8), 126.9 (C7), 127.4 (C12), 132.2 (C13), 135.6 (C11), 147.8 (C5), 155.5 (C10), 167.9 (C(3)O); **HRMS:** (ESI^+) Calculated for $C_{96}H_{120}N_8O_{28}Na$: 1856.8138, found $[M+Na]^+$: 1856.8145; Calculated for $C_{96}H_{120}N_8O_{28}Na_2$: 939.9015, found $[M+2Na]^{2+}$: 939.9019. **Retention time:** (C18, 70:30 water/acetone → 100% acetone over 70 mins, 16 mL/min flow rate) 60 mins (analytical), 59 mins (preparative).

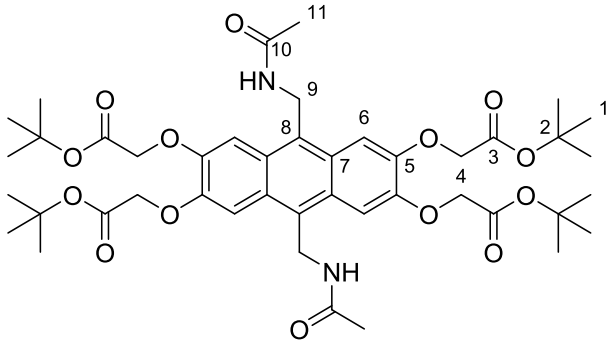
1,1'-(1,2-phenylene)bis(3-benzylurea) (63)[121]



Under an inert N_2 atmosphere, benzyl isocyanate (0.5 g, 3.76 mmol) was added to a solution of phenylene-1,2-diamine **38** (81 mg, 0.75 mmol) in dry CH_2Cl_2 (30 mL) and stirred at reflux for 16 hours. The resultant precipitate was filtered and purified by flash column chromatography (100% CH_2Cl_2) to afford **63** (0.25 g, 0.68 mmol, 90%) as a white solid**. ^{1}H NMR:** (400 MHz, $(CD_3)_2SO$): δ 4.26 (d, *J* = 5.8

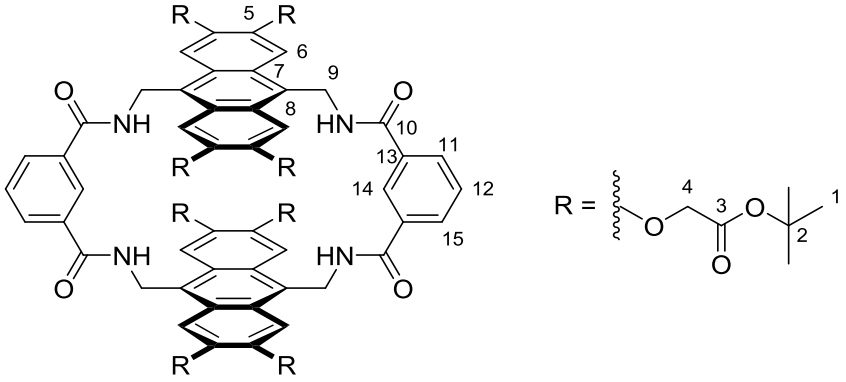
Hz, 4H, C(5)H$_2$), 6.92-6.99 (m, 4H, N(11)H and C(1)H)), 7.16-7.31 (m, 10H, Aryl-H), 7.49 (dd, *J* = 6, 3.6 Hz, 2H, C(15)H), 7.87 (s, 2H, N(10)H), [lit.[121] (400 MHz, $(CD_3)_2SO$) δ 4.28, 7.30, 6.96, 7.55, 7.90]; **^{13}C NMR:** (100 MHz, $(CD_3)_2SO$): δ 43.42 (C(5)H$_2$), 123.8 (C(1)H), 124.0 (C(2)H), 127.1, 127.7, 128.7 (Aryl-CH), 132.0 (C3), 140.7 (Aryl-C6), 156.4 (C5), [lit.[121] (100 MHz, $(CD_3)_2SO$) δ 42.9, 123.2, 123.4, 126.6, 127.2, 128.2, 140.2, 155.9]; **LRMS:** (ESI$^+$) Calculated for $C_{22}H_{22}N_4O_2Na$: 397.2, found $[M+Na]^+$: 397.2.

tetra-*tert*-butyl-2,2',2'',2'''-((9,10-bis(acetamidomethyl)anthracene-2,3,6,7-tetrayl)tetrakis(oxy)) tetraacetate (64)



Under an inert N_2 atmosphere, **41** (30 mg, 0.04 mmol) was dissolved in anhydrous degassed CH_2Cl_2 (10 mL). DIPEA was added (0.02 mL, 0.12 mmol) and then acetyl chloride (0.04 mL, 0.48 mmol) at 0 °C and then the mixture stirred at room temperature for 16 hours. MeOH (2 mL) was added and the solvent removed under vacuum. The crude product was dissolved in CH_2Cl_2 (30 mL), washed with water (30 mL), 5% $NaHCO_3$ (30 mL), dried ($MgSO_4$) and the solvent removed under vacuum. The crude product was then purified by flash column chromatography (2% MeOH:CH_2Cl_2) to yield **64** (28 mg, 0.033 mmol, 84%) as a pale brown solid. **^{1}H NMR:** (400 MHz, $(CDCl_3)$): δ 1.49 (s, 36H, 3 × C(1)H$_3$), 1.85 (s, 6H, C(11)H$_3$), 4.79 (s, 8H, C(4)H$_2$), 5.07 (s, 4H, 2 × C(9)H$_2$), 7.48 (s, 4H, 4 × C(6)H); **^{13}C NMR:** (100 MHz, $(CDCl_3)$): δ 23.3 (C(11)H$_3$), 28.2 (C(1)H$_3$), 36.4 (C(9)H$_2$), 66.7 (C(4)H$_2$), 82.1 (C(2)$(CH_3)_3$), 105.2 (C(6)H), 125.8 (C8), 126.1 (C7), 148.2 (C5), 167.7 (C(3)O), 169.4 (C(10)O); **LRMS:** (ESI$^+$) Calculated for $C_{44}H_{60}N_2O_{14}Na$: 863.4, found $[M+Na]^+$: 863.4.

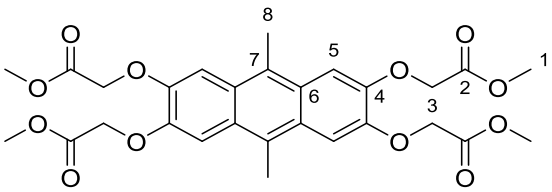
***tert*-butyl protected tetra-amide macrocycle (67)**



Under an inert N_2 atmosphere, **42** (65 mg, 0.086 mmol) was dissolved in dry degassed dichloromethane (600 mL) and DIPEA (1 mL) was added. Pentaflurophenyl ester **66** (43 mg, 0.086 mmol) in dry degassed dichloromethane (45 mL) was added *via* syringe pump (1 mL/hour) and the reaction stirred at room temperature for 3 days. The solvent was removed under vacuum and the crude product purified by reverse phase HPLC (acetone/water) and then freeze dried to afford **67** (36 mg, 0.021 mmol, 24%) as a white solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 1.46 (s, 72H, C(1)$\underline{H}_3$), 4.77 (s, 16H, C(4)$\underline{H}_2$), 5.46 (d, *J* = 5.1 Hz, 8H, C(9)$\underline{H}_2$), 7.48 (t, *J* = 7.9 Hz, 2H, C(12)$\underline{H}$), 7.58 (s, 8H, C(6)$\underline{H}$), 8.01 (d, *J* = 8.1 Hz, 2H, C(11)$\underline{H}$, 8.13 (d, *J* = 7.6 Hz, 2H, C(15)$\underline{H}$), 8.30 (s, 2H, C(14)$\underline{H}$); **HRMS:** (ESI^+) Calculated for $C_{96}H_{116}N_4O_{28}Na$: 1796.7702, found $[M+Na]^+$: 1796.7711. **Retention time:** (40:60 water/acetone → 100% acetone over 40 mins, 16 mL/min flow rate) 13 mins (analytical), 12 mins (preparative).

*No ^{13}C data was obtained due to instability of compound **67** during prolonged characterisation.*

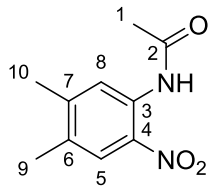
tetramethyl 2,2',2'',2'''-((9,10-dimethylanthracene-2,3,6,7-tetrayl)tetrakis(oxy))tetraacetate (68)



Under an inert N_2 atmosphere, **45** (0.5 g, 1.85 mmol) was dissolved in dry THF (100 mL). K_2CO_3 (1.15 g, 8.32 mmol) and methyl bromoacetate (1.1 mL, 11.1 mmol) were added and the reaction mixture stirred at reflux for 24 hours. The mixture was cooled to room temperature and the solvent removed under vacuum. The crude solid was then dissolved in CH_2Cl_2 (100 mL), washed with water (40 mL), brine (40 mL) and dried ($MgSO_4$). The solvent was removed under vacuum and the crude residue purified by flash column chromatography (1% MeOH:CH_2Cl_2) to yield **68** (400 mg, 0.72 mmol, 39%) as an orange brown solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 2.79 (s, 6H, C(8)$\underline{H}_3$), 3.84 (s, 12H, C(1)$\underline{H}_3$), 4.87 (s, 8H, C(3)$\underline{H}_2$), 7.36 (s, 4H, C(5)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 14.6 ($\underline{C}$8), 52.3 ($\underline{C}$1), 66.3 ($\underline{C}$3), 106.2

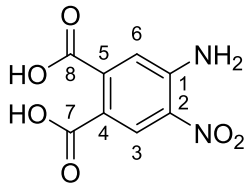
(C5), 124.8 (C7), 126.3 (C6), 147.1 (C4), 169.2 (C2); ν_{max} 2955, 1754, 1495, 1435, 1380, 1214, 1188, 1068 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{28}H_{30}O_{12}Na$: 581.1629, found $[M+Na]^+$: 581.1623.

N-(4,5-dimethyl-2-nitrophenyl)acetamide (77)[61]



4,5-dimethyl-2-nitroaniline **78** (2 g, 12 mmol) was suspended in glacial acetic acid (24 mL) and heated to 90 °C. Acetic anhydride (1.2 mL, 13 mmol) was added and the mixture stirred at reflux for 2 hours. The reaction mixture was cooled to room temperature and poured into water (300 mL). The yellow precipitate was filtered, washed with water and recrystallized from ethanol to yield **77** (2.4 g, 11.6 mmol, 97%) as a yellow crystalline solid. **m.p.** (ethanol) 105-107 °C [lit.[122] (ethanol) 101-104 °C]; **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 2.25 (s, 6H, C(9, 10)H$_3$), 2.32 (s, 3H, C(1)H$_3$), 7.93(s, 1H, C(8)H), 8.51 (s, 1H, C(5)H), 10.26 (br s, 1H, NH), [lit.[61] (400 MHz, $CDCl_3$) δ 2.27, 2.28, 2.35, 7.98, 8.54, 10.30]; **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 19.1 (C10), 20.5 (C9), 25.6 (C1), 122.6 (C8), 125.9 (C5), 132.3 (C7), 132.7 (C6), 134.1 (C3), 146.8 (C4), 168.9 (C(2)O); ν_{max} 3341, 2987, 2901, 1708, 1695, 1576, 1323, 1151, 759 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{10}H_{12}N_2O_3Na$: 231.0740, found $[M+Na]^+$: 231.0745.

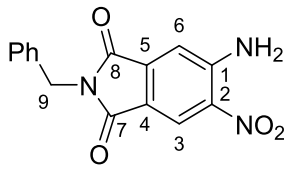
4-amino-5-nitrophthalic acid (76)[61]



Under an inert N_2 atmosphere, **77** (2 g, 9.6 mmol) and $KMnO_4$ (6 g, 37.9 mmol)) was suspended in water (50 mL) and stirred at reflux for 3 days. Additional $KMnO_4$ (3g, 19 mmol) was added halfway through the reaction time. The resultant brown precipitate was filtered hot and washed with water. The yellow filtrate was acidified to pH 3 with 1M HCl, extracted with EtOAc (3 × 100 mL), washed with brine (100 mL) and dried ($MgSO_4$). The solvent was removed under vacuum to yield **76** (0.77 g, 2.88 mmol, 60%) as a yellow orange crystalline solid**. m.p.** (H_2O) 232-236 °C [lit.[123] (H_2O) 238 °C); **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 7.25 (s, 1H, C(6)H), 7.63 (br s, 2H, NH$_2$), 8.70 (s, 1H, C(3)H), 11.50 (br s, 2H, C(7, 8)O$_2$H) [lit.[61] (400 MHz, $(CD_3)_2SO$) δ 7.91, 8.26, 13.0]; **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 117.7 (C4), 118.7 (C6), 129.3 (C3), 132.6 (C5), 141.5 (C2), 146.9 (C1), 167.8 (C7), 169.0 (C8); ν_{max} 3486, 3364, 2972,

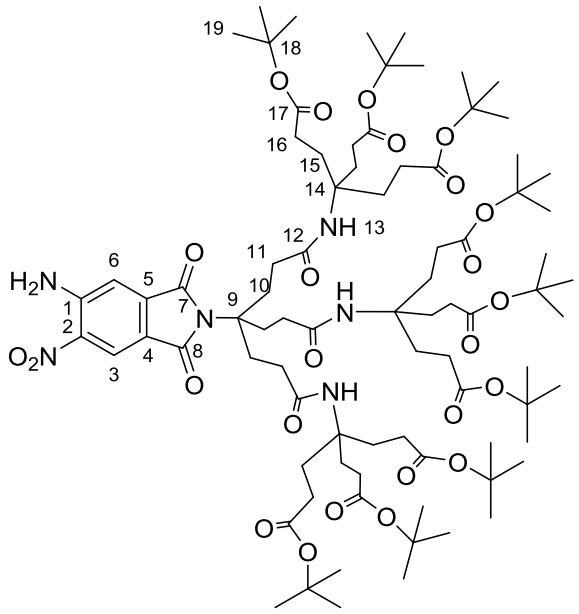
2901, 1712, 1681, 1626, 1502, 1252, 1057, 882 cm^{-1}; **LRMS:** (EI) Calculated for $C_8H_6N_2O_6$: 226.0 Found $[M]^+$: 226.1.

5-amino-2-benzyl-6-nitroisoindoline-1,3-dione (82)



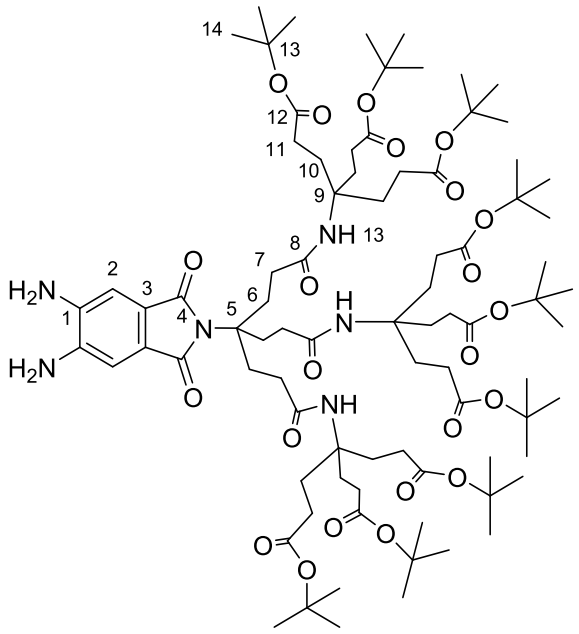
Under an inert N_2 atmosphere, **76** (40 mg, 0.15 mmol) was dissolved in anhydrous THF (8 mL) and acetic anhydride (0.4 mL, 2.5 mmol) added. The reaction mixture was stirred at reflux for 4 hours and then cooled to room temperature. The solvent was removed under vacuum and the crude residue co-evaporated with toluene (3 × 20 mL) to afford a yellow solid **75** which was used without further purification. This crude product was then dissolved in anhydrous toluene (5 mL) under an N_2 atmosphere and benzylamine (19 µL, 0.16 mmol) was added. The reaction mixture was stirred a reflux for 16 hours, after which the solvent was removed under vacuum and the crude solid purified by flash column chromatography (1% MeOH:CH_2Cl_2) to afford **82** (37 mg, 0.13 mmol, 84%) as a yellow solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): 4.81 (s, 2H, C(9)H_2), 7.21-7.39 (m, 5H, Aryl-H), 7.57 (s, 1H, C(6)H), 7.99 (br s, 2H, NH_2), 8.46 (s, 1H, C(3)H); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 41.3 (C9), 114.4 (C6), 117.2 (C4), 122.2 (C3), 122.3, 127.9, 128.4 (Aryl-CH), 133.6 (C5), 136.8 (Aryl-C) 137.3 (C2), 150.2 (C1), 165.9 (C7), 165.9 (C8); **ν_{max}** 3466, 3353, 3051, 2971, 1766, 1697, 1642, 1598, 1530, 1260, 1070, 743 cm^{-1}; **LRMS:** (EI) Calculated for $C_{15}H_{11}N_3O_4$: 297.1, found $[M]^+$: 297.1.

G2MM 2-nitro-4,5-phthalimido aniline (84)



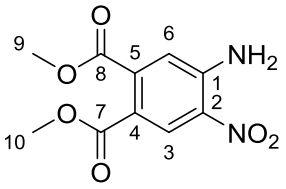
Under an inert N_2 atmosphere, **76** (40 mg, 0.15 mmol) was dissolved in anhydrous THF (5 mL) and acetic anhydride (0.4 mL, 2.5 mmol) added. The reaction mixture was stirred at reflux for 4 hours and then cooled to room temperature. The solvent was removed under vacuum and the crude residue co-evaporated with toluene (3 × 20 mL) to afford a yellow solid **75** which was used without further purification. This crude product was then dissolved in anhydrous toluene (5 mL) under an N_2 atmosphere and G2MM amine **83** (140 mg, 0.1 mmol) was added. The reaction mixture was stirred a reflux for 16 hours, after which the solvent was removed under vacuum and the crude solid purified by flash column chromatography (1% MeOH:CH_2Cl_2) to afford **84** (104 mg, 0.064 mmol, 64%) as a yellow solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): 1.42 (s, 81H, C(19)$\underline{H}_3$), 1.86-.195 (m, 24H, C(10, 15)$\underline{H}_2$), 2.12-2.22 (m, 24, C(11, 16)$\underline{H}$), 6.81 (br s, 3H, N(13)$\underline{H}$), 6.81 (s, 1H, C(6)$\underline{H}$), 8.74 (s, 1H, C(3)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 28.2 (C19), 29.7, 29.9 (C16/15), 31.3, 31.5 (C10/11), 55.1 (C14), 57.7 (C9), 80.8 (C18), 117.3 (C4), 118.2 (C6), 129.8 (C5), 130.5 (C3), 145.4 (C2), 147.9 (C1), 166.6 (C7), 172.8 (C12), 173.0 (C17); **ν_{max}** 3464, 3348, 2987, 2901, 1725, 1681, 1628, 1542, 1393, 1250, 1056 cm^{-1}; **LRMS:** (ESI) Calculated for $C_{84}H_{136}N_6O_{25}Na$: 1629.0. found $[M+Na]^+$: 1629.0.

G2MM 1,2-diamino-4,5-phthalimide (88)



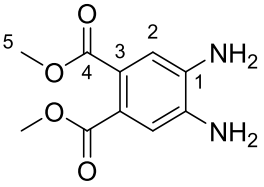
Under an inert N_2 atmosphere, a solution of **84** (200 mg, 0.12 mmol) in MeOH (10 mL) was added to Pd/C (10 mg). The reaction vessel was then purged with hydrogen (1 atm) and the reaction mixture stirred at room temperature for 3 hours. After which, the reaction mixture was filtered through celite and washed with CH_2Cl_2, and the filtrate concentrated under vacuum. The crude product was then purified by flash column chromatography (4% MeOH:CH_2Cl_2) to afford **88** (179 mg, 0.11 mmol, 90%) as light brown solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 1.41 (s, 81H, C(14)$\underline{H}_3$), 1.95 (m, 18H, C(11)$\underline{H}_2$), 2.12 (m, 6H, C(6)$\underline{H}_2$), 2.19 (m, 18H, C(10)$\underline{H}_2$), 2.27 (m, 6H, C(7)$\underline{H}_2$), 7.07 (s, 2H, C(2)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 28.1 ($\underline{C}$14), 29.9 ($\underline{C}$10), 30.0 ($\underline{C}$11), 31.9 ($\underline{C}$7), 32.3 ($\underline{C}$6), 53.5 ($\underline{C}$5), 57.5 ($\underline{C}$9), 80.7 ($\underline{C}$13) 110.6 ($\underline{C}$2), 124.8 ($\underline{C}$3), 139.6 ($\underline{C}$1), 169.4 ($\underline{C}$4); **HRMS:** (ESI^+) Calculated for $C_{84}H_{138}N_6O_{23}Na$: 1621.9706, found $[M+Na]^+$: 1621.9709.

Dimethyl 4-amino-5-nitrophthalate (89)



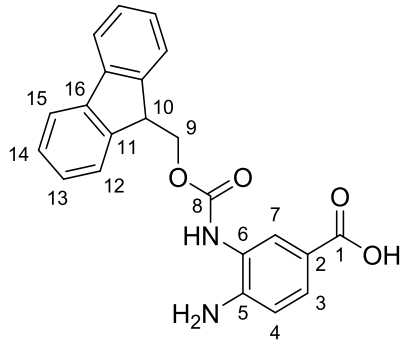
76 (0.8 g, 3.5 mmol) was dissolved in MeOH (30 mL) and concentrated H_2SO_4 (0.5 mL) added. The reaction mixture was stirred at reflux for 3 hours and then the solvent was removed under vacuum. The residue was dissolved in EtOAc (60 mL), washed with 5% $NaHCO_3$ (60 mL), brine (60 mL) and dried ($MgSO_4$). The solvent was removed under vacuum and the crude solid purified by flash column chromatography (100% CH_2Cl_2) to yield **89** (0.72g, 2.8 mmol, 80%) as an orange solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 3.88 (s, 3H, C(10)H$_3$), 3.92 (s, 3H, C(9)H$_3$), 6.91 (s, 1H, C(6)H), 7.26 (br s, 2H, NH$_2$), 8.74 (s, 1H, C(3)H); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 52.5 (C9), 53.1 (C10), 116.9 (C4), 118.3 (C6), 129.7 (C3), 131.2 (C5), 140.7 (C2), 146.6 (C1), 164.7 (C7), 168.0 (C8); **ν_{max}** 3486, 3342, 2987, 2901, 1736, 1697, 1621, 1502, 1435, 1339, 1250, 1027, 762 cm^{-1}; **LRMS:** (ESI$^+$) Calculated for $C_{10}H_{10}N_2O_6Na$: 277.0, found [M+Na]$^+$: 277.1.

Dimethyl 4,5-diamino-phthalate (90)[124]



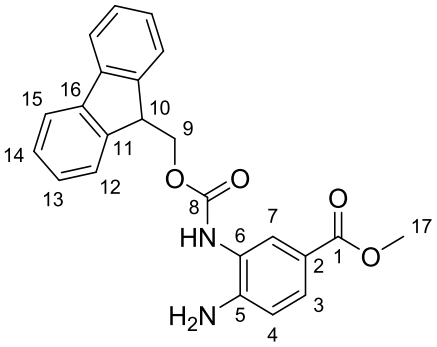
Under an inert N_2 atmosphere, a solution of **89** (100 mg, 0.39 mmol) in MeOH (10 mL) was added to Pd/C (10 mg). The reaction vessel was then purged with hydrogen (1 atm) and the reaction mixture stirred at room temperature for 1 hour. After which, the reaction mixture was filtered through celite and washed with CH_2Cl_2, and the filtrate concentrated under vacuum. The crude product was then purified by flash column chromatography (5% MeOH:CH_2Cl_2) to afford **90** (81 mg, 0.36 mmol, 93%) as light brown solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 3.84 (s, 6H, C(5)H$_3$), 7.02 (s, 2H, C(2)H), [lit.[124] (500 MHz, $CDCl_3$) δ 3.83, 6.99]; **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 52.4 (C5), 116.6 (C2), 124.1 (C3), 1136.6 (C1), 168.4 (C4), [lit.[124] (125 MHz, $CDCl_3$) δ 52.3, 116.5, 123.9, 136.7, 168.4]; **ν_{max}** 3363, 2997, 2950, 1694, 1624, 1574, 1522, 1434, 1323, 1262, 1209, 1128, 1050 cm^{-1}; **HRMS:** (ESI$^+$) Calculated for $C_{10}H_{12}N_2O_4Na$: 247.0689, found [M+Na]$^+$: 247.0685.

3-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-4-aminobenzoic acid (100)[67]



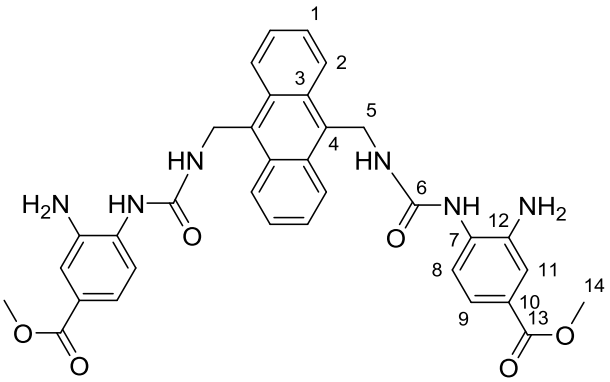
To a suspension of 3,4-diaminobenzoic acid **99** (1 g, 6.60 mmol), in MeCN/0.1M aqueous $NaHCO_3$ (1:1, 40 mL), was added Fmoc-Su (2.20 g, 6.60 mmol) over 30 minutes. The mixture was rapidly stirred at room temperature for 24 hours and then acidified to pH 1 with 1M HCl. The pink precipitate was then filtered and washed with cold Et_2O (20 mL), hexane (20 mL) and methanol (20 mL) before being air dried to afford **99** (2.14 g, 5.94 mmol, 90%) as a light pink solid. **^{1}H NMR:** (400 MHz, $((CD_3)_2SO)$: δ 4.25-4.34 (m, 1H, C(10)H), 4.35-4.45 (m, 2H, C(9)H_2), 5.68 (s, 2H, NH_2), 6.71 (d, *J* = 8.4 Hz, 1H, C(4)H), 7.30-7.38 (m, 2H, C(13)H), 7.40-7.46 (m, 2H, C(14)H), 7.50 (d, *J* = 7.8 Hz, 1H, C(3)H), 7.69-7.78 (m, 2H, C(15)H), 7.79 (s, 1H, C(7)H), 7.91 (d, *J* = 7.5 Hz, 2H, C(12)H), 8.77 (s, 1H, C(O)NH), 12.10 (s, 1H, CO_2H), [lit.[67] (400 MHz, $(CD_3)_2SO$) δ 4.29-4.40, 6.72, 7.33-7.36, 7.43, 7.51, 7.74-7.80, 7.90, 8.78]; **^{13}C NMR:** (100 MHz, $((CD_3)_2SO)$: δ 41.8 (C10), 66.4 (C9), 114.7 (C4), 118.0 (C5), 120.6 (C15), 122.2 (C6), 125.8 (C13), 127.6 (C12), 128.3 (C3, 7), 128.2 (C14), 141.2 (C16), 144.2 (C11), 155.0 (C8), 167.7 (C1), [lit.[67] (100 MHz, $(CD_3)_2SO$) δ 46.6, 114.1, 117.4, 120.0, 125.2, 127.0, 127.6, 140.6, 143.7, 154.4, 167.1]; **ν_{max}** 3409, 3336, 3218, 3066, 2932, 1682, 1605, 1576, 1415, 1341, 1297, 1221, 1034, 736 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{22}H_{19}N_2O_4$: 375.1339, found $[M+H]^+$: 375.1359, Calculated for $C_{22}H_{18}N_2O_4Na$: 397.1159, found $[M+Na]^+$: 397.1168.

methyl 3-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-4-aminobenzoate (107)



To a suspension of methyl 3,4-diaminobenzoate **106** (1 g, 6.0 mmol), in MeCN/0.1M aqueous $NaHCO_3$ (1:1, 30 mL), was added Fmoc-Su (2 g, 6.0 mmol) over 30 minutes. The mixture was rapidly stirred at room temperature for 24 hours. The light orange precipitate was then filtered and washed with cold hexane (20 mL) before being air dried to afford **107** (2.38 g, 5.99 mmol, 99%) as a pale orange solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 3.85 (s, 3H, C(17)H$_3$), 4.12-4.28 (m, 3H, C(9, 10)H), 6.35 (br s, 2H, NH$_2$), 6.72 (d, *J* = 8.4 Hz, 1H, C(4)H), 7.29-7.35 (m, 2H, C(13)H), 7.40 (t, *J* = 7.8 Hz, 2H, C(14)H), 7.49-7.64 (m, 2H, C(12)H), 7.76 (d, *J* = 7.7 Hz, 3H, C(3, 15)H), 7.83 (d, *J* = 1.9 Hz, 1H, C(7)H); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 47.2 (C10), 51.8 (C17), 67.1 (C9), 115.9 (C4), 120.0 (C15), 121.5 (C6), 125.0 (C13), 127.1 (C12), 127.8 (C7), 128.2 (C14), 129.5 (C3) 141.4 (C16), 143.6 (C11), 155.2 (C8), 166.7 (C1); **ν_{max}** 3346, 3068, 2936, 1688, 1612, 1582, 1425, 1347, 1294, 1230, 1044, 745 **HRMS:** (ESI^+) Calculated for $C_{23}H_{21}N_2O_4$: 389.1496, found $[M+H]^+$: 389.1489, calculated for $C_{23}H_{20}N_2O_4Na$: 411.1315, found $[M+Na]^+$: 411.1311.

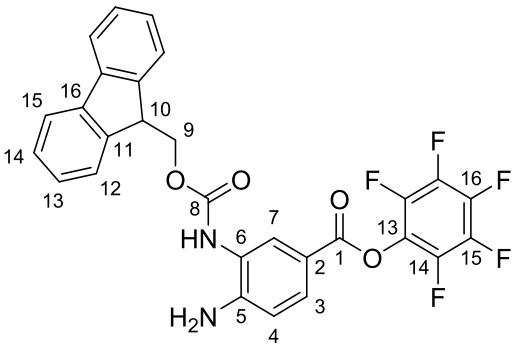
Diamino methyl ester anthracene half receptor (109)



Under an inert N_2, **107** (754 mg, 1.94 mmol) was dissolved in anhydrous dichloromethane (50 mL) and **60** (280 mg, 0.97 mmol) was added. The reaction was heated to reflux for 16 hours and the solvent removed under vacuum. The crude residue was suspended in cold methanol and filtered to afford **108** (312 mg, 50.4 mmol, 52%) as a crude yellow solid. Under an inert N_2 atmosphere, **108** was dissolved in piperidine (20% in DMF, 20 mL) and stirred at room temperature for 4 hours. The piperidine was co-

evaproated with toluene and the crude residue triturated in dichloromethane. The precipitate was filtered and air dried to afford **109** (90 mg, 0.15 mmol, 90%) as a pale yellow solid. **^{1}H NMR:** (400 MHz, ($(CD_3)_2SO$): δ 3.76 (s, 6H, C(14)H$_3$), 5.11 (s, 4H, NH$_2$), 5.34 (d, *J* = 5.3 Hz, 4H, C(5)H$_2$), 7.18 (dd, *J* = 2.0 and 8.4 Hz, 2H, C(9)H), 7.27 (d, *J* = 2.1 Hz, 1H, C(11)H), 7.44 (t, *J* = 5.4 Hz, 2H, NH), 7.63 (dd, *J* = 3.3 and 6.9 Hz, 4H, C(1)H), 7.96 (d, *J* = 8.5 Hz, 1H, C(8)H), 8.53 (dd, *J* = 3.3 and 6.9 Hz, 4H, C(2)H), 8.58 (s, 2H, NH); **^{13}C NMR:** (100 MHz, ($(CD_3)_2CO$): δ 35.9 (C5), 52.0 (C14), 116.1 (C9), 118.7 (C11), 119.3 (C8), 125.6 (C2), 126.4 (C10), 126.5 (C1), 127.4 (C3), 129.4 (C4), 131.4 (C12), 138.1 (C7), 155.4 (C6), 166.4 (C13); **HRMS:** (ESI^+) Calculated for $C_{34}H_{32}N_6O_6Na$: 643.2276, found $[M+Na]^+$: 643.2256.

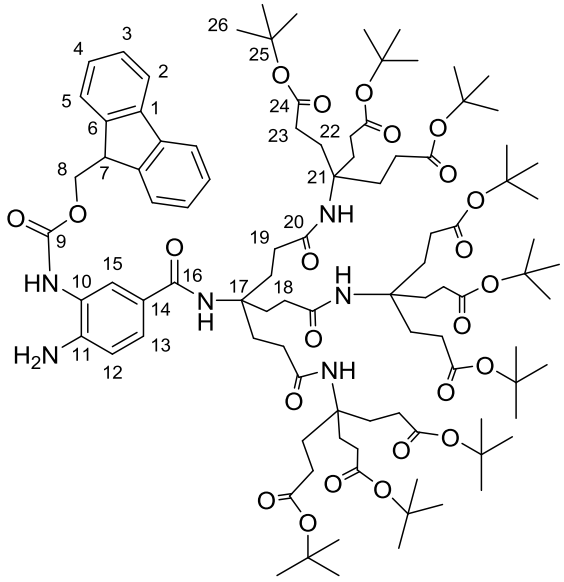
Fmoc protected pentafluorophenyl ester (112)



Under an inert N_2 atmosphere, **100** (200 mg, 0.53 mmol) was suspended in anhydrous THF (20 ml). A solution of pentafluorophenol (140mg) and DCC (160 mg) in anhydrous THF (10 mL) was added, and the reaction heated to 55 °C upon which the precipitate disappeared. The reaction was stirred for 16 hours, cooled to room temperature and concentrated under vacuum. The crude residue was suspended in dichloromethane and filtered, and the filtrate concentrated under vacuum. The crude product was then purified by flash column chromatography (1:1, 40/60 → 1:4, 40/60 petroleum ether:EtOAc) to afford **112** (240 mg, 0.45 mmol, 84%) as a white solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 4.24 (s, 1H, C(10)H), 4.59 (s, 2H, C(9)H$_2$), 6.22 (br s, 2H, NH$_2$), 6.79 (d, *J* = 8.6 Hz, 1H, C(4)H), 7.31 (s, 2H, C(13)H), 7.41 (t, *J* = 7.5 Hz, 2H, C(14)H), 7.58 (br s, 2H, C(12c)H), 7.77 (d, *J* = 7.5 Hz, 2H, C(15)H), 7.91 (d, *J* = 8.6 Hz, 1H, C(4)H), 7.95 (s, 1H, C(7)H); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 47.2 (C10), 67.2 (C9), 116.0 (C3), 121.6 (C6), 127.0 (C7), 127.1 (C12), 129.5 (C3), 131.0 (C4), 143.5 (C11), 155.2 (C8), 161.9 (C1); **^{19}F NMR:** (377 MHz, ($CDCl_3$)): δ -162,68 (t, *J* = 21.1 Hz, 2F, C(15)F), -158.57 (t, *J* = 21.1 Hz, 1F, C(16)F), -152.57 (m, 2F, C(14)F), **HRMS:** (ESI^+) Calculated for $C_{28}H_{17}F_5N_2O_4Na$: 563.1001, found $[M+Na]^+$: 563.0981.

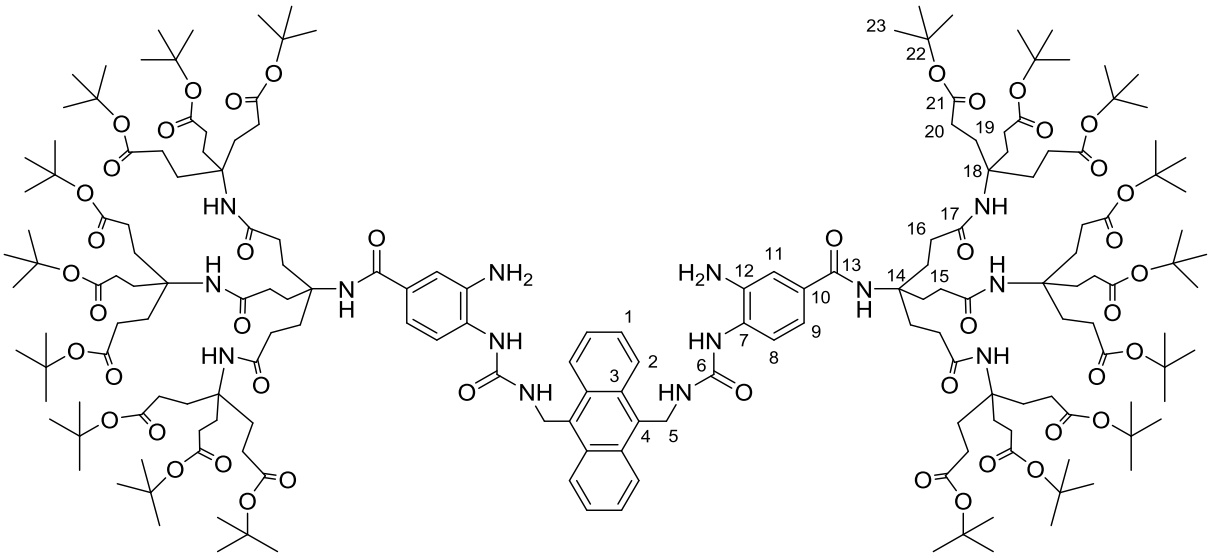
The ^{13}C signals for fluorinated carbons 13, 14, 15 and 16 were of too weak intensity to be observed.

Fmoc protected *tert*-butyl G2MM linker (113)



Under an inert N_2 atmosphere, **100** (100 mg, 0.267 mmol), HBTU (100 mg, 0.267 mmol) and HOBt (36 mg, 0.267 mmol) were suspended in anhydrous THF (10 mL). DIPEA (0.15 mL, 0.534 mmol) was added and the reaction stirred at room temperature for 10 minutes. G2MM amine **83** (192 mg, 0.133 mmol) was then added and the reaction stirred for 24 hours. The solvent was then removed under vacuum and the crude residue purified by flash column chromatography (5% MeOH:CH_2Cl_2) to afford **84** (229 mg, 0.127 mmol, 96%) as an off white solid. **^{1}H NMR**: (400 MHz, ($CDCl_3$): δ 1.41 (s, 81H, C(26)H_3), 1.95 (m, 18H, C(23)H_2), 2.11 (t, *J* = 7.2 Hz, 6H, C(18)H_2), 2.18 (m, 18H, C(22)H_2), 2.26 (t, *J* = 7.2 Hz, 6H, C(19)H_2), 4.14-4.34 (br m, 3H, C(7)H and NH_2), 4.46 (d, *J* = 7.4 Hz, 2H, C(8)H_2), 6.11 (s, 3H, NH), 6.77 (d, *J* = 8.4 Hz, 1H, C(13)H), 7.26-7.31 (m, 2H, C(4)H), 7.39 (t, *J* = 7.4 Hz, 2H, C(3)H), 7.57-7.69 (m, 2H, C(5)H), 7.72 (d, *J* = 8.7 Hz, 1H, C(2)H), 7.76 (d, *J* = 7.6 Hz, 3H, C(2)H), 7.78 (d, *J* = 2.1 Hz, 1H, C(15)H), 8.59 (s, 1H, NH); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 28.0 (C26), 29.8 (C22), 29.9 (C23), 31.8 (C19), 32.2 (C18), 47.2 (C7), 53.4 (C17), 57.4 (C21), 67.3 (C8), 80.6 (C25) 116.6 (C12), 119.9 (C2), 122.6 (C10), 124.6 (C14), 125.3 (C4), 126.0 (C15), 126.8 (C13), 127.0 (C5), 127.6 (C3), 141.3 (C1), 143.8 (C6), 145.3 (C11), 154.9 (C9), 166.6 (C16), 172.7 (C24), 173.1 (C20); **ν_{max}** 2977, 2963, 1752, 1723, 1689, 1637, 1535, 1367, 1242, 1151, 1098, 844 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{98}H_{150}N_6O_{24}Na_2$: 921.0260, found $[M+2Na]^{2+}$: 921.0252. **Retention time:** (C18, 20:80 water/acetone → 100% acetone over 30 mins, 16 mL/min flow rate) 13 mins (analytical), 13 mins (preparative).

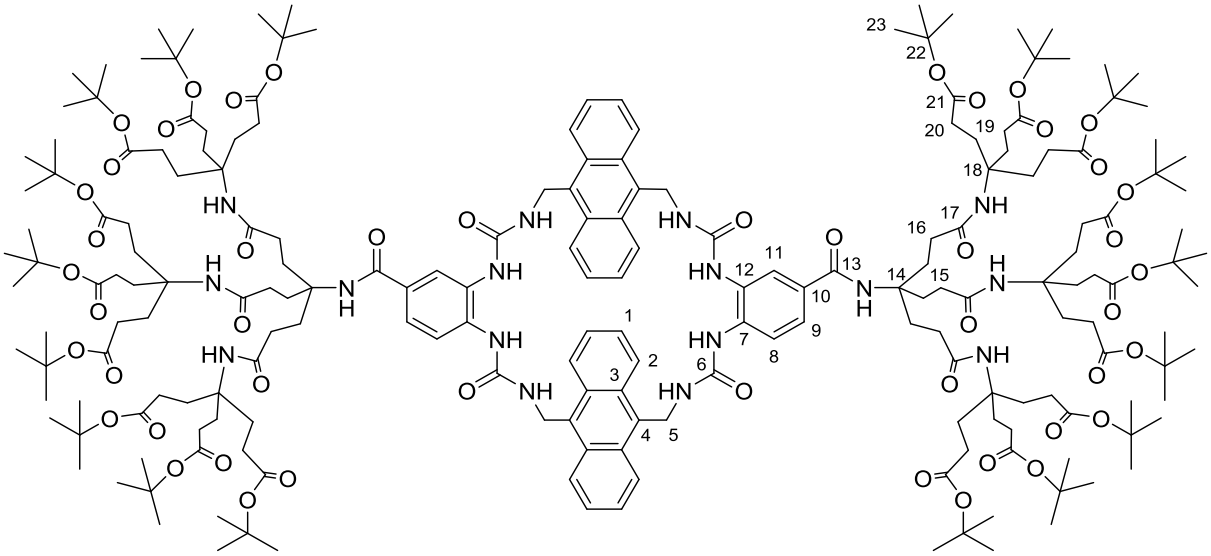
Diamino *tert*-butyl protected G2MM anthracene half receptor (119)



Under an inert N_2 atmosphere, **113** (500 mg, 0.288 mmol) and **60** (37 mg, 0.130 mmol) were dissolved in anhydrous dichloromethane (8 mL). Pyridine (60 µL, 0.74 mmol) was added and the reaction heated to reflux for 16 hours. The reaction was cooled to room temperature and the solvent removed under vacuum. The crude residue was purified by flash column chromatography (4% MeOH:CH_2Cl_2) to afford the Fmoc protected product **118** (464 mg, 0.12 mmol, 92%) as a white solid. Conversion to **118** was confirmed by limited NMR studies* and high resolution mass spectrometry (ESI^+): Calculated for $C_{214}H_{312}N_{14}O_{50}Na_2$: 1963.1077, found $[M+2Na]^{2+}$: 1963.1067. Under an inert N_2 atmosphere, **118** (260 mg, 0.066 mmol) was dissolved in anhydrous dichloromethane (10 mL) and cooled to 0 °C. DBU (45 µL, 0.28 mmol) was added and the reaction mixture warmed to room temperature and stirred for 2 hours. The solvent was removed under vacuum and the crude product purified by flash column chromatography (5% MeOH:CH_2Cl_2) to afford **119** (215 mg, 0.063 mmol, 95%) as an off white solid. **^{1}H NMR:** (400 MHz, (CD_3OD): δ 1.43 (s, 162H, C(23)$\underline{H}_3$), 1.93 (m, 36H, C(20)$\underline{H}_2$), 2.08 (m, 12H, C(15)$\underline{H}_2$), 2.18 (m, 48H, C(19, 16)$\underline{H}_2$), 5.36 (s, 4H, C(5)$\underline{H}_2$), 7.18 (dd, *J* = 2.1, 8.3 Hz, 2H, C(9)$\underline{H}$), 7.26 (d, *J* = 2.1 Hz, 2H, C(11)$\underline{H}$), 7.41 (d, *J* = 8.3 Hz, 2H, C(8)$\underline{H}$), 7.41 (s, 6H, N$\underline{H}$), 7.60 (dd, *J* = 3.3, 6.9 Hz, 4H, C(1)$\underline{H}$), 7.89 (s, 2H, N$\underline{H}$), 8.47 (dd, *J* = 3.3, 6.9 Hz, 4H, C(2)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 27.1 (C23), 29.1 (C20), 29.3 (C21), 31.0 (C15), 31.1 (C16), 57.4 (C18), 58.0 (C14), 80.2 (C22), 115.9 (C11), 117.4 (C9) 123.1 (C8), 124.6 (C2), 125.9 (C1), 128.5 (C10), 130.1 (C4), 130.5 (C3), 131.4 (C7), 140.0 (C12), 156.7 (C6), 168.6 (C13), 173.0 (C21), 174.1 (C17); **ν_{max}** 3321, 2976, 2938, 1725, 1653, 1534, 1455, 1391, 1308, 1249, 1148, 1101, 954, 844 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{184}H_{293}N_{14}O_{46}Na$: 1730.5524, found $[M+H+Na]^{2+}$: 1730.5507.

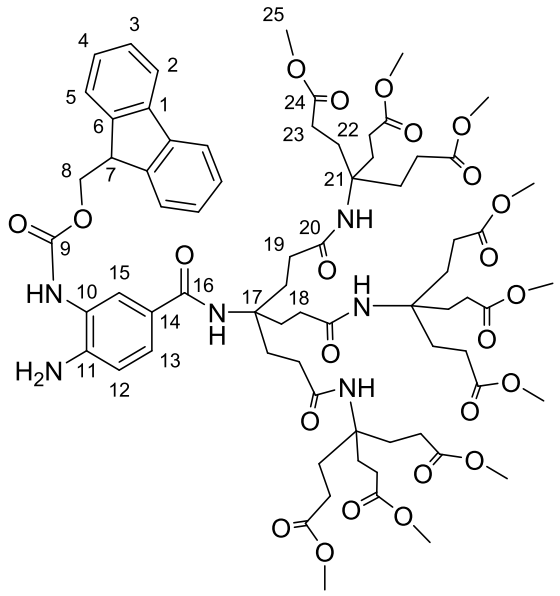
* *Limited NMR studies were only possible due to believed slow conformational exchange of **118** resulting in very broad signals of low intensity.*

***tert*-butyl protected G2MM anthracene tetra urea macrocycle (120)**



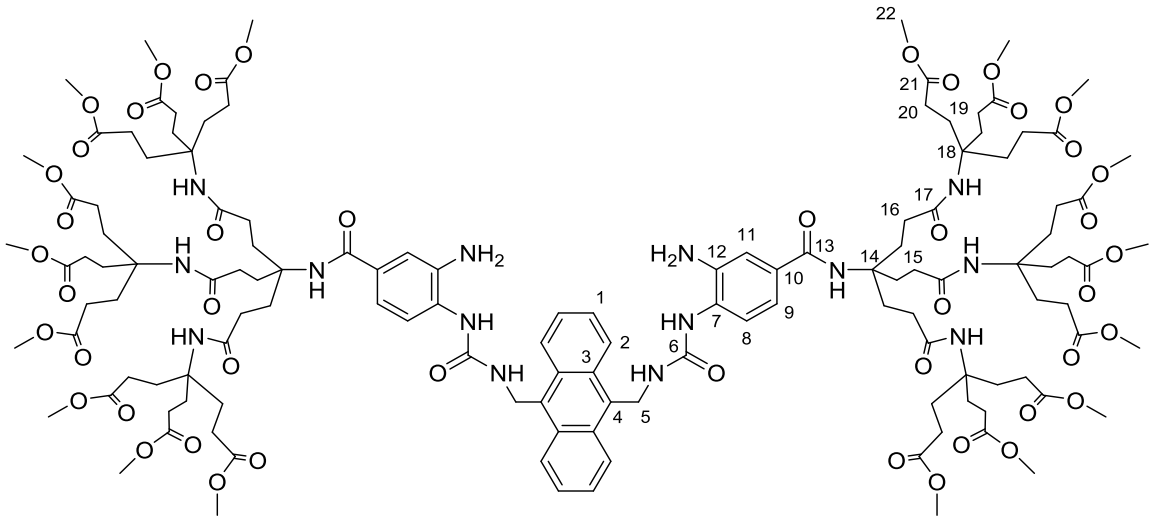
Under an inert N_2 atmosphere, **119** (63 mg, 0.029 mmol) and DMAP (7 mg, 0.058 mmol) were dissolved in anhydrous degassed dichloromethane (60 mL) and heated to reflux. **60** (8.4 mg, 0.029 mmol) in anhydrous degassed dichloromethane (5 mL) was added and the reaction stirred at reflux for 2 days. The solvent was then removed under vacuum and the crude product purified by reverse phase HPLC (acetone/water) and then freeze dried to afford **120** (41 mg, 0.011 mmol, 38%) as a white solid. **^{1}H NMR:** (400 MHz, (CD_3OD): δ 1.45 (s, 162H, C(23)$\underline{H}_3$), 1.99 (m, 36H, C(20)$\underline{H}_2$), 2.17 (m, 12H, C(15)$\underline{H}_2$), 2.24 (m, 48H, C(19)$\underline{H}_2$), 2.31 (m, 12H, C(16)$\underline{H}_2$), 5.39 (s, 8H, C(5)$\underline{H}_2$), 7.31, 7.44 (br s, 4H, C(1)$\underline{H}$), 7.49 (s, 6H, N$\underline{H}$), 7.70 (dd, *J* = 2.1, 8.5 Hz, 2H, C(9)$\underline{H}$), 7.89 (d, *J* = 8.5 Hz, 2H, C(8)$\underline{H}$), 7.98 (s, 2H, C(11)$\underline{H}$), 8.39 (br s, 8H, C(2)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 27.0 ($\underline{C}$23), 29.1 ($\underline{C}$20), 29.3 ($\underline{C}$21), 30.8 ($\underline{C}$15), 31.1 ($\underline{C}$16), 57.3 ($\underline{C}$18), 58.1 ($\underline{C}$14), 80.3 ($\underline{C}$22), 121.6 ($\underline{C}$9), 121.77 ($\underline{C}$12) 124.4 ($\underline{C}$2), 124.7 ($\underline{C}$11), 125.8, 125.9 ($\underline{C}$1), 130.0 ($\underline{C}$10), 130.4 ($\underline{C}$3), 130.6 ($\underline{C}$3), 131.4 ($\underline{C}$7, 12), 156.0, 156.8 ($\underline{C}$6), 168.1 ($\underline{C}$13), 173.1 ($\underline{C}$21), 174.1, 174.2 ($\underline{C}$17); **HRMS:** (ESI^+) Calculated for $C_{202}H_{304}N_{16}O_{48}Na_3$: 1264.3861, found $[M+3Na]^{3+}$: 1264.3835, Calculated for $C_{202}H_{304}N_{16}O_{48}Na_4$: 954.0369, found $[M+4Na]^{4+}$: 954.0378. **Retention time:** (C18, 20:80 water/acetone → 100% acetone over 30 mins, 16 mL/min flow rate) 23 mins (analytical), 21 mins (preparative).

Fmoc protected methyl ester G2MM linker (127)



113 (215 mg, 0.12 mmol) was dissolved in dichloromethane (5 mL) and TFA (5 mL) added dropwise. The solution was stirred at room temperature for 16 hours and the TFA and solvent evaporated under a flow of N_2. The residue was redissolved in methanol (5 mL) and trimethyl orthoformate (5 mL). HCl (5% v/v, 0.5 mL) was added and the mixture stirred for 24 hours. The solvent was then removed under vacuum and the crude product purified by flash column chromatography (5% MeOH:CH_2Cl_2) to afford **127** (147 mg, 0.10 mmol, 87%). **^{1}H NMR:** (400 MHz, ($CDCl_3$): δ 1.99 (m, 18H, C(23)$\underline{H}_2$), 2.108(m, 6H, C(18)$\underline{H}_2$), 2.25 (m, 24H, C(22, 19)$\underline{H}_2$), 3.60 (s, 27H, C(25)$\underline{H}_3$), 4.23 (m, 1H, C(7)$\underline{H}$), 4.48 (m, 2H, C(8)$\underline{H}_2$), 6.31 (s, 3H, N$\underline{H}$), 6.72 (d, *J* = 8.4 Hz, 1H, C(13)$\underline{H}$), 7.26 (m, 3H, C(4, 12)$\underline{H}$), 7.38 (t, *J* = 7.4 Hz, 2H, C(3)$\underline{H}$), 7.63 (m, 2H, C(5)$\underline{H}$), 7.75 (d, *J* = 7.9 Hz, 3H, C(2, 15)$\underline{H}$), 8.12 (s, 1H, N$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): 28.3 (C$\underline{22}$), 29.6 (C$\underline{23}$), 31.8 (C$\underline{19}$), 32.2 (C$\underline{18}$), 47.2 (C$\underline{7}$), 51.8 (C$\underline{25}$), 57.3 (C$\underline{21}$), 58.2 (C$\underline{17}$), 67.1 (C$\underline{8}$), 116.4 (C$\underline{12}$), 120.0 (C$\underline{2}$), 122.6 (C$\underline{10}$), 124.4 (C$\underline{14}$), 125.1 (C$\underline{4}$), 125.2 (C$\underline{15}$), 126.6 (C$\underline{13}$), 127.7 (C$\underline{5}$), 127.8 (C$\underline{3}$), 141.3 (C$\underline{1}$), 143.6 (C$\underline{11}$), 143.8 (C$\underline{6}$), 155.0 (C$\underline{9}$), 166.6 (C$\underline{16}$), 173.3 (C$\underline{20}$), 173.8 (C$\underline{24}$); **ν_{max}** 3330, 2976, 2961, 1727, 1658, 1531, 1452, 1367, 1249, 1150, 846 cm^{-1}**; HRMS:** (ESI$^+$) Calculated for $C_{71}H_{96}N_6O_{24}Na$: 1439.6363, found $[M+Na]^+$: 1439.6353.

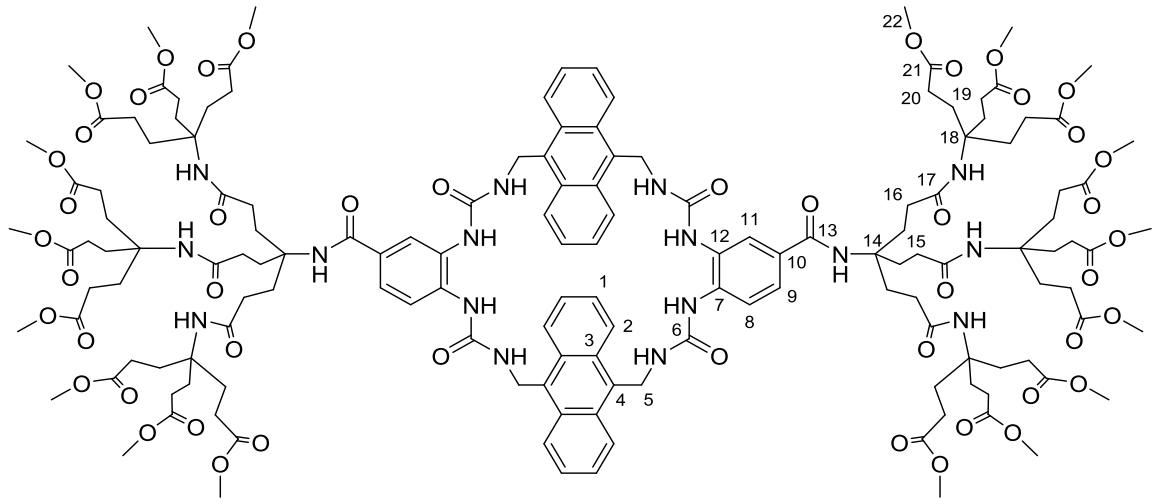
Diamino methyl ester protected G2MM anthracene half receptor (129)



Under an inert N_2 atmosphere, **127** (500 mg, 0.353 mmol) and **60** (42 mg, 0.147 mmol) were dissolved in anhydrous dichloromethane (5 mL). Pyridine (50 µL, 0.62 mmol) was added and the reaction heated to reflux for 16 hours. The reaction was cooled to room temperature and the solvent removed under vacuum. The crude residue was purified by flash column chromatography (4% → 6% MeOH:CH_2Cl_2) to afford the Fmoc protected product **128** (410 mg, 0.131 mmol, 89%) as a white solid. Conversion to **128** was confirmed by limited NMR studies* and high resolution mass spectrometry (ESI^+): m/z calculated for $C_{180}H_{204}N_{14}O_{50}Na_2$: 1584.1834, found $[M+2Na]^{2+}$: 1584.1849. Under an inert N_2 atmosphere, **128** (450 mg, 0.144) was dissolved in anhydrous dichloromethane (10 mL) and cooled to 0 °C. DBU (85 µL, 0.58 mmol) was added and the reaction mixture warmed to room temperature and stirred for 4 hours. The solvent was removed under vacuum and the crude product purified by flash column chromatography (5% MeOH:CH_2Cl_2) to afford **129** (335 mg, 0.125 mmol, 87%) as an off white solid. **^{1}H NMR:** (400 MHz, (CD_3OD): δ 1.94 (m, 36H, C(20)$\underline{H}_2$), 2.07 (m, 12H, C(15)$\underline{H}_2$), 2.24 (m, 48H, C(19, 16)$\underline{H}_2$), 3.60 (s, 54H, C(22)$\underline{H}_3$), 4.94 (s, 4H, C(5)$\underline{H}_2$), 7.05 (d, *J* = 8.3 Hz, 2H, C(9)$\underline{H}$), 7.15 (m, 4H, C(11, 8)$\underline{H}$), 7.37 (s, 6H, N$\underline{H}$), 7.48 (dd, *J* = 3.3, 6.9 Hz, 4H, C(1)$\underline{H}$), 7.85 (s, 2H, N$\underline{H}$), 8.21 (dd, *J* = 3.3, 6.9 Hz, 4H, C(2)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 28.5 (C20), 29.3 (C21), 31.0 (C15), 31.1 (C16), 51.8 (C22), 57.3 (C18), 58.1 (C14), 116.1(C11), 117.5 (C9) 123.0 (C8), 124.4 (C2), 125.8 (C1), 128.6 (C10), 130.0 (C4), 130.5 (C3), 131.4 (C7), 140.1 (C12), 155.5 (C6), 166.7 (C13), 173.3 (C17), 173.9 (C21); **HRMS:** (ESI^+) Calculated for $C_{130}H_{184}N_{14}O_{46}Na_3$: 915.7400, found $[M+3Na]^{3+}$: 915.7415. **Retention time:** (C18, water/acetone, 50:50 → 100% acetone over 30 mins, 16 mL/min flow rate) 19 mins (analytical), 19 mins (preparative).

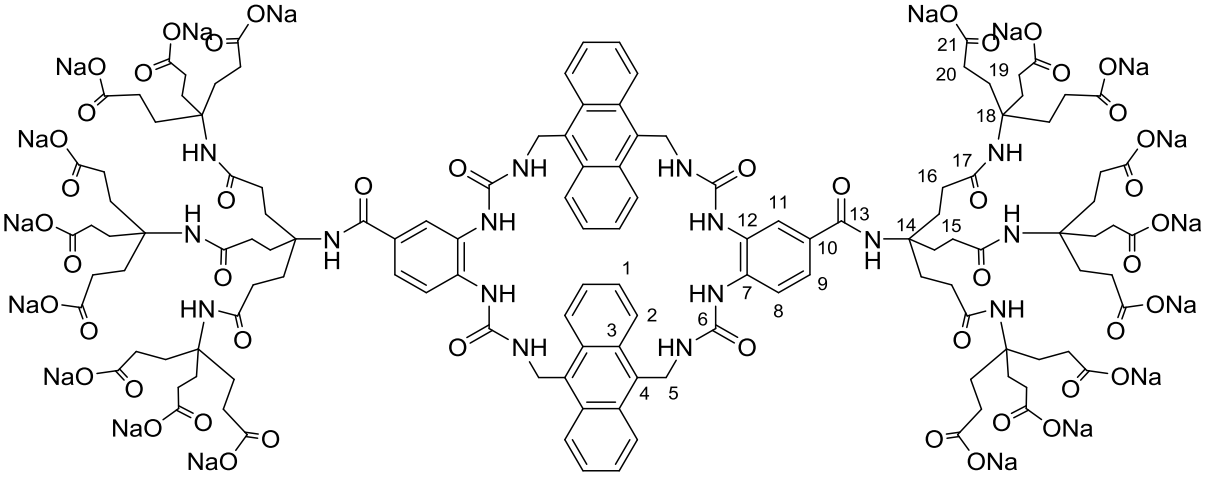
* *Limited NMR studies were only possible due to believed slow conformational exchange of **128** resulting in very broad signals of low intensity.*

Methyl ester protected G2MM anthracene tetra urea macrocycle (126)



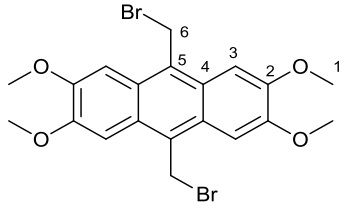
Under an inert N_2 atmosphere, **129** (100 mg, 0.037 mmol) and DMAP (9 mg, 0.074) were dissolved in anhydrous degassed dichloromethane (32 mL) and heated to reflux. **60** (10.7 mg, 0.037 mmol) in anhydrous degassed dichloromethane (5 mL) was added and stirred at reflux for 2 days. The solvent was then removed under vacuum and the crude product purified by reverse phase HPLC (acetone/water) and then freeze dried to afford **126** (37 mg, 0.013 mmol, 34%) as a pale yellow solid. **^{1}H NMR:** (400 MHz, $(CD_3)_2SO$): δ 1.89 (m, 36H, C(20)$\underline{H}_2$), 1.95 (m, 12H, C(15)$\underline{H}_2$),), 2.12 (m, 12H, C(16)$\underline{H}_2$), 2.22 (m, 36H, C(19)$\underline{H}_2$), 5.25 (s, 8H, C(5)$\underline{H}_2$), 7.32 (s, 6H, N$\underline{H}$), 7.42 (br s, 4H, C(1)$\underline{H}$), 7.52 (m, 6H, C(1, 9)$\underline{H}$), 7.74 (s, 2H, N$\underline{H}$), 7.87 (d, *J* = 8.5 Hz, 2H, C(8)$\underline{H}$), 8.07 (s, 2H, C(11)$\underline{H}$), 8.37 (br s, 8H, C(2)$\underline{H}$); **^{13}C NMR:** (100 MHz, $(CD_3)_2SO$): δ 28.3 (C$\underline{20}$), 29.2 (C$\underline{19}$), 29.5, 29.8 (C$\underline{15}$), 30.7, 31.0 (C$\underline{16}$), 35.7 (C$\underline{5}$), 51.8 (C$\underline{22}$), 56.8 (C$\underline{18}$), 57.9 (C$\underline{14}$), 120.0 (C$\underline{8}$), 122.8 (C$\underline{9}$), 123.3 (C$\underline{11}$), 125.4 (C$\underline{2}$), 125.7 (C$\underline{12}$), 126.4, 126.5 (C$\underline{1}$), 129.0, 129.5 (C$\underline{3}$), 129.9 (C$\underline{4}$), 135.6 (C$\underline{7}$) 155.1, 155.7 (C$\underline{6}$), 166.0 (C$\underline{13}$), 172.9 (C$\underline{17}$), 173.7 (C$\underline{21}$); **HRMS:** (ESI^+) Calculated for $C_{148}H_{196}N_{16}O_{48}Na_2$: 1506.1603, found $[M+2Na]^{2+}$: 1506.1621, Calculated for $C_{148}H_{196}N_{16}O_{48}Na_3$: 1011.7699, found $[M+3Na]^{3+}$: 1011.7687. **Retention time:** (C18, water/acetone, 50:50 → 100% acetone over 30 mins, 16 mL/min flow rate) 17 mins (analytical), 17 mins (preparative).

Deprotected G2MM anthracene tetra urea macrocycle (1)



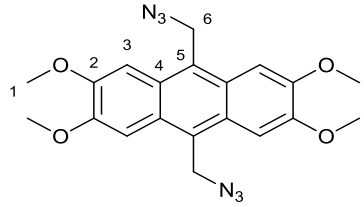
Protected receptor **120** (30 mg, 0.008 mmol) was dissolved in dichloromethane (HPLC grade, 2.7 ml) and cooled to 0 °C. Trifluoroacetic acid (TFA) (0.3 mL) was added dropwise and the reaction warmed to room temperature and stirred for 4 hours. The solvent was then removed under a flow of nitrogen, then the residue was co-evaporated with toluene (3 x 10 mL) to remove residual TFA, suspended in water and freeze dried. The product was then purified by preparative HPLC eluting with 100% Water (buffered with 0.1 % TFA) → 100% methanol over 40 minutes. The solvent was removed under vacuum, the residue co-evaporated with toluene (3 × 10 mL), the product suspended in water and freeze dried. The solid was then suspended in water, neutralised to pH 7.4 with NaOH (aq), filtered and then freeze dried to afford **1** as a white solid (21 mg, 0.0068 mmol, 85%). **^{1}H NMR:** (400 MHz, 75 °C, D_2O): δ 2.40-2.50 (m, 36H, C(19)$\underline{H}_2$), 2.58-2.71 (m, 48H, C(20)$\underline{H}_2$, C(15)$\underline{H}_2$), 2.80-2.90 (m, 12H, C(16)$\underline{H}_2$), 5.63, 5.81 (br s, 4H, C(5)$\underline{H}_2$), 7.43-7.50, 7.98-8.06 (br m, 4H, C(1)$\underline{H}$),8.25 (br s, 2H, C(11)$\underline{H}$), 8.27 (d, 2H, C(9)$\underline{H}$), 8.39 (d, 2H, C(8)$\underline{H}$), 7.63 (d, *J* = 8.3 Hz, 3H, C(10)$\underline{H}$), 8.58-8.65, 8.82-8.89 (br m, 4H, C(2)$\underline{H}$). **^{13}C NMR:** (100 MHz, 75 °C, (D_2O): δ 30.4 ($\underline{C}$19), 30.6 ($\underline{C}$15), 30.9 ($\underline{C}$16), 31.2 ($\underline{C}$20), 37.7, 37.9 ($\underline{C}$5), 57.4 ($\underline{C}$18), 58.3 ($\underline{C}$14), 124.6, 124.8 ($\underline{C}$2), 124.9 ($\underline{C}$11, $\underline{C}$9), 125.7, 125.9 ($\underline{C}$1), 128.3 ($\underline{C}$12), 129.8 ($\underline{C}$4), 130.1 ($\underline{C}$10), 130.5, 130.7 ($\underline{C}$3), 135.6 ($\underline{C}$7), 156.4, 157.2 ($\underline{C}$6), 168.4 ($\underline{C}$13), 174.8 ($\underline{C}$17), 182.0 ($\underline{C}$21). **Retention time:** (C18, 100% Water (buffered with 0.1 % TFA) → 100% methanol over 40 minutes, 16 mL/min flow rate) 28 mins (analytical), 28 mins (preparative).

2,3,6,7-Tetramethoxy-9,10-bis(bromomethyl)anthracene, (136)



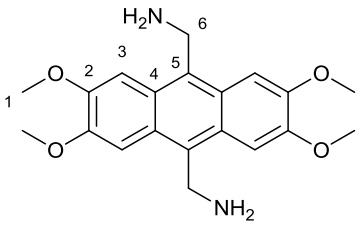
Under an inert N_2 atmosphere, **91** (2 g, 6.1 mmol), NBS (4 g, 22.6 mmol) and ABCN (73 mg, 0.3 mmol) were dissolved in anhydrous dichloromethane (150 mL), and the mixture stirred at reflux for 4 hours. The mixture was then cooled to 0 °C and filtered. The solid was then dried under high vacuum to afford **136** (2.1 g, 4.3 mmol, 71%) as a bright yellow solid. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.11 (s, 12H, C(1)H), 5.35 (s, 4H, C(6)H), 7.40 (s, 4H, C(3)H), **^{13}C NMR** (100 MHz, $CDCl_3$) 28.7 (C6), 56.0 (C1), 101.8 (C3), 125.7 (C5), 125.9 (C4), 150.1 (C2); $\boldsymbol{\nu}_{max}$ 3002, 2942, 2829, 1633, 1535, 1498, 1458, 1430, 1371, 1247, 1201, 1160, 1147, 1030, 830 cm^{-1}; **HRMS** $(ESI)^+$ Calculated for $C_{20}H_{20}Br_2O_4Na$: 506.9600, 504.9621, found $[M+Na]^+$: 506.9611, 504.9635.

2,3,6,7-Tetramethoxy-9,10-bis(azidoomethyl)anthracene, (137)



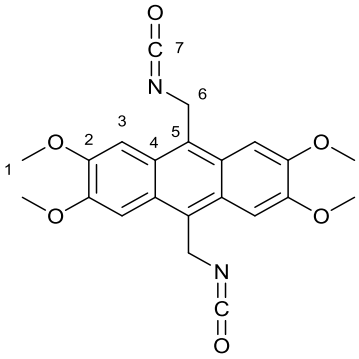
Under an inert N_2 atmosphere, **136** (2.7 g, 5.58 mmol) and NaN_3 were suspended in anhydrous MeCN (70 mL). The mixture was stirred at reflux for 16 hours and then cooled to room temperature and the solvent was evaporated *in vacuo*. The remaining residue was then suspended in water (200 mL) and filtered. The solid was washed with ethanol (3 × 100 mL) and dried under high vacuum to afford **137** (1.6 g, 3.9 mmol, 70%) as a brown solid. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.10 (s, 12H, C(1)H), 5.19 (s, 4H, C(6)H), 7.42 (s, 4H, C(3)H), **^{13}C NMR** (100 MHz, $CDCl_3$) 47.7 (C6), 56.1 (C1), 101.8 (C3), 123.8 (C5), 126.9 (C4), 150.2 (C2); $\boldsymbol{\nu}_{max}$ 2997, 2968, 2936, 2831, 2105, 1632, 1535, 1497, 1455, 1432, 1371, 1244, 1200, 1190, 1165, 1032, 833 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{20}H_{20}N_6O_4Na$: 431.1438, found $[M+Na]^+$ 431.1448.

2,3,6,7-Tetramethoxy-9,10-bis(aminomethyl)anthracene, (138)



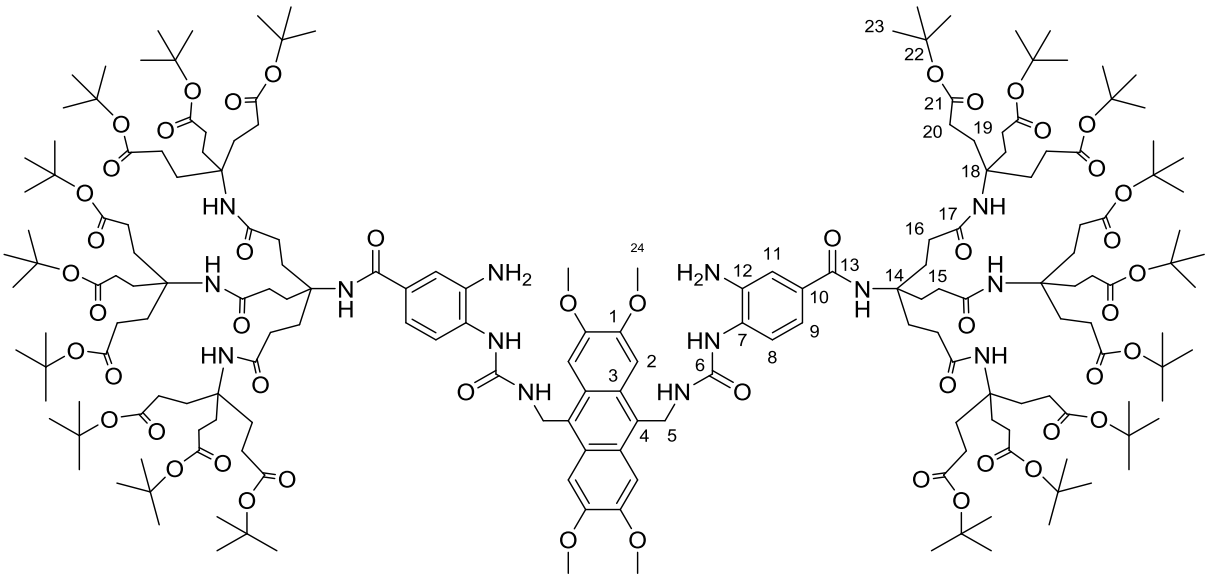
Under an inert N_2 atmosphere, **137** (1.6 g, 3.90 mmol) and PPh_3 (8 g, 31.4 mmol) were suspended in degassed THF (80 mL). Degassed water (4 mL) was added and the reaction heated to 60 °C for 16 hours. The reaction was cooled to room temperature and the solvent removed under vacuum. The crude residue was suspended in toluene (200 mL), filtered, washed with toluene (2 × 100 mL) and dried under high vacuum to afford **138** (1.05 g, 2.96 mmol, 76%) as a pale brown solid. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.08 (s, 12H, C(1)H), 4.69 (s, 4H, C(6)H), 7.49 (s, 4H, C(3)H), **^{13}C NMR** (100 MHz, $CDCl_3$) δ 39.5 (C6), 56.0 (C1), 101.9 (C3), 125.3 (C5), 125.5 (C4), 149.7 (C2); **ν_{max}** 3361, 2998, 2910, 2828, 1633, 1535, 1494, 1455, 1433, 1241, 1205, 1168, 1042, 1029, 842 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{20}H_{24}N_2O_4Na$: 379.1628, found $[M+Na]^+$ 379.1626.

9,10-bis(isocyanatomethyl)-2,3,6,7-tetramethoxyanthracene (139)



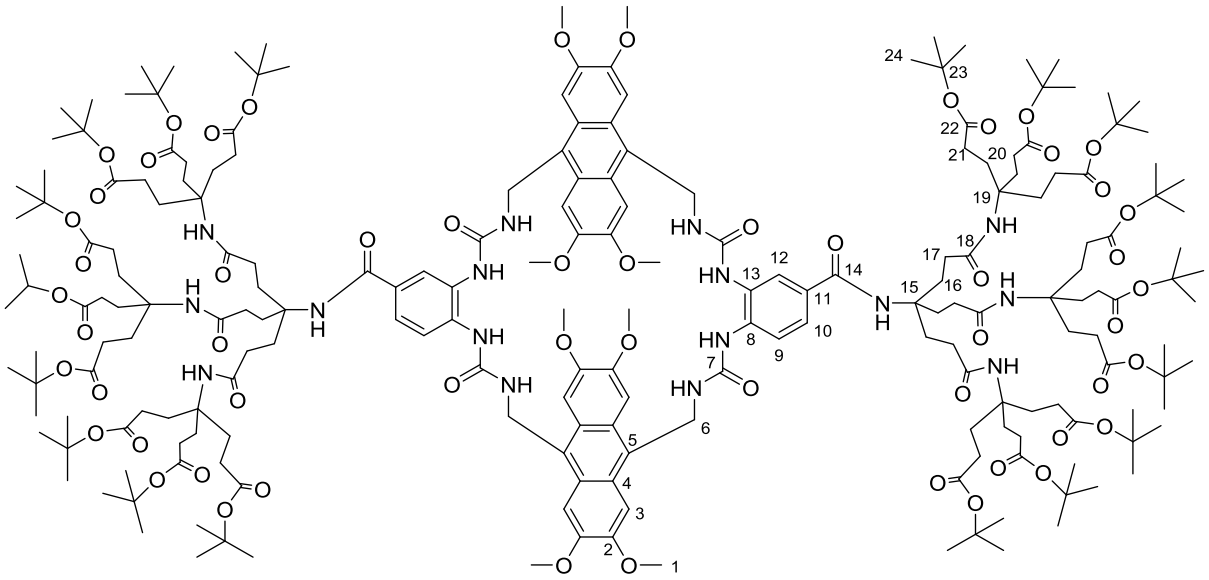
Under an inert N2 atmosphere, a flask was charged with triphosgene (324 mg, 1.1 mmol) and anhydrous toluene (15 mL) was added. A suspension of **138** (200 mg, 0.55 mmol) in anhydrous toluene (5 mL) was added dropwise and the reaction mixture stirred at reflux for 2 hours. The reaction mixture was cooled and the solvent removed under high vacuum. The crude solid was resuspended in dichloromethane (50 mL) and filtered. The filtrate was collected and the solvent removed under vacuum to afford **139** (121 mg, 0.30 mmol, 54%) as a brown solid. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.11 (s, 12H, C(1)H), 5.21 (s, 4H, C(6)H), 7.37 (s, 4H, C(3)H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 40.1 (C6), 56.0 (C1), 101.3 (C3), 124.5 (C7), 125.4 (C5), 125.7 (C4), 150.2 (C2); **ν_{max}** 2934, 2832, 2255, 1498, 1435, 1245, 1204, 1169, 1028 cm^{-1}; **HRMS:** (ESI^+) calculated for $C_{22}H_{20}N_2O_6Na$: 431.1214, found $[M+Na]^+$ 431.1218.

Diamino *tert*-butyl protected methoxy-anthracene half receptor (141)



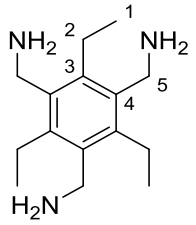
Under an inert N_2 atmosphere, **113** (507 mg, 0.283 mmol) and **139** (50 mg, 0.13 mmol) were dissolved in anhydrous dichloromethane (8 mL). Pyridine (60 µL, 0.74 mmol) was added and the reaction heated to reflux for 16 hours. The reaction was cooled to room temperature and the solvent removed under vacuum. The crude residue was purified by reverse phase HPLC (acetone/water) to afford **140** (336 mg, 0.84 mmol, 65%) as a white solid. Conversion to **140** was confirmed by limited NMR studies* and high resolution mass spectrometry (ESI^+): m/z calculated for $C_{218}H_{320}N_{14}O_{54}Na_2$: 2023.1288, found 2023.1299. Under an inert N_2 atmosphere, **140** (100 mg, 0.028 mmol) was dissolved in anhydrous dichloromethane (10 mL) and cooled to 0 °C. DBU (50 µL, 0.31 mmol) was added and the reaction mixture warmed to room temperature and stirred for 1 hour. The solvent was removed under vacuum and the crude product purified by flash column chromatography (6% MeOH:CH_2Cl_2) to afford **141** (91 mg, 0.026 mmol, 92%) as an off white solid. **^{1}H NMR:** (400 MHz, (CD_3OD): δ 1.42 (s, 162H, C(23)H$_3$), 1.84-2.0 (m, 36H, C(20)H$_2$), 2.02-2.14 (m, 12H, C(15)H$_2$), 2.13-2.31 (m, 48H, C(19, 16)H$_2$), 3.95 (br s, 12H, C(24)H$_3$), 4.58 (br s, 4H, C(5)H$_2$), 7.13 (d, *J* = 8.3 Hz, 2H, C(9)H), 7.23 (s,2H, C(11)H), 7.31 (d, *J* = 8.3 Hz, 2H, C(8)H), 7.38 (br s, 4H, C(2)H), 7.41 (s, 6H, NH), 7.89 (s, 2H, NH); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 27.1 (C23), 29.1 (C20), 29.3 (C21), 30.8 (C15), 31.1 (C16), 54.9 (C24), 57.4 (C18), 58.1 (C14), 80.3 (C22), 101.9 (C2), 115.8 (C11), 117.3 (C9) 123.3 (C8), 126.1 (C4), 126.1 (C3), 128.2 (C10), 131.6 (C7), 140.3 (C12), 149.5 (C1), 157.0 (C6), 168.5 (C13), 173.0 (C21), 174.1 (C17); **HRMS:** (ESI^+) Calculated for $C_{188}H_{300}N_{14}O_{50}Na_3$: 1208.3702, found $[M+3Na]^{2+}$:1208.3682. **Retention time:** (C18, 20:80 water/acetone → 100% acetone over 30 mins, 16 mL/min flow rate) 22 mins (analytical), 20 mins (preparative).* *Limited NMR studies were only possible due to believed slow conformational exchange of **140** resulting in very broad signals of low intensity.*

***t*ert-butyl protected octa-methoxy anthracene tetra urea macrocycle (97)**



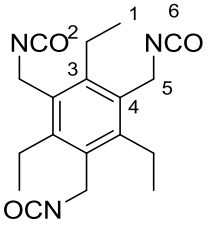
Under an inert N_2 atmosphere, **96** (25 mg, 0.007 mmol) and DMAP (1.7 mg, 0.014 mmol) were dissolved in anhydrous degassed dichloromethane (12 mL) and heated to reflux. **94** (2.7 mg, 0.007 mmol) in anhydrous degassed dichloromethane (2 mL) was added and the reaction stirred at reflux for 2 days. The solvent was then removed under vacuum and the crude product purified by reverse phase HPLC (acetone/water) and then freeze dried to afford **89** (Xmg, X mmol, X%) as a white solid. **^{1}H NMR:** (400 MHz, 75 °C, D_2O): δ 2.40-2.50 (m, 36H, C(19)$\underline{H}_2$), 2.58-2.71 (m, 48H, C(20)$\underline{H}_2$, C(15)$\underline{H}_2$), 2.80-2.90 (m, 12H, C(16)$\underline{H}_2$), 5.63, 5.81 (br s, 4H, C(5)$\underline{H}_2$), 7.43-7.50, 7.98-8.06 (br m, 4H, C(1)$\underline{H}$),8.25 (br s, 2H, C(11)$\underline{H}$), 8.27 (d, 2H, C(9)$\underline{H}$), 8.39 (d, 2H, C(8)$\underline{H}$), 7.63 (d, *J* = 8.3 Hz, 3H, C(10)$\underline{H}$), 8.58-8.65, 8.82-8.89 (br m, 4H, C(2)$\underline{H}$). **^{13}C NMR:** (100 MHz, (D_2O): δ 30.4 (C19), 30.6 (C21), 30.9 (C16), 31.2 (C17), 36.1, 36.4 (C6), 54.8, 54.9 (C1), 57.3, (C15), 58.2 (C19), 80.2 (C23), 101.5, 101.9 (C3), 121.1 (C9), 124.3, (C10, C12), 124.9 (C11, C9), 125.7 (C5), 125.9 (C4), 128.1 (C13), 129.6 (C11), 135.7 (C8), 149.4, 149.5 (C2), 156.0, 157.0 (C7), 168.0 (C14), 173.0 (C18), 174.2 (C22); **HRMS:** (ESI^+) Found $[M+3Na]^{3+}$:1339.7453. **Retention time:** (C18, 20:80 water/acetone → 100% acetone over 30 mins, 16 mL/min flow rate) 21 mins (analytical), 19 mins (preparative).

1,3,5-triethyl-2,4,6-tris(aminomethyl)benzene (154)[125]



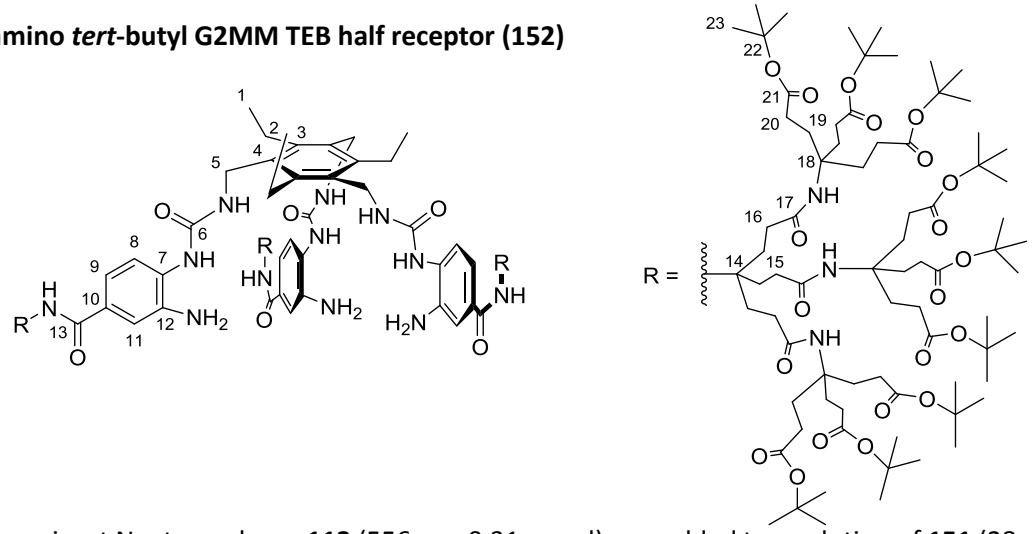
Under an inert N_2 atmosphere, 1,3,5-tris(bromomethyl)-2,4,6-triethylbenzene **155** (324 mg, 0.74 mmol) was dissolved in anhydrous DMF (4.5 mL) and NaN_3 (157 mg, 2.42 mmol) added. The reaction was heated to 60 °C for 16 hours. The reaction mixture as then diluted with ethyl acetate (20 mL) and washed with water (3 × 20 mL), dried ($MgSO_4$) and filtered. DMF (4 mL) was added to the filtrate and the solvent removed under vacuum to a volume of ~4 mL. Conversion to tris-azide **156** was confirmed by ^{1}H NMR (220 mg, 0.68 mmol, 92%). The resultant DMF solution was transferred to a degassed anhydrous solution of THF (22 mL) and PMe_3 (1M in THF, 4.1 mL) under an inert N_2 atmosphere. The reaction mixture was stirred at room temperature for 1 hour and degassed H_2O (5 mL) added, with the reaction mixture stirred for a further 16 hours. The solvent and excess PMe_3 was then evaporated by bubbling N_2 through the solution, and the crude residue suspended in H_2O (~10 mL). The suspension was then freeze dried to afford **154** (148 mg, 0.61 mmol, 90%) as a white solid. **^{1}H NMR:** (400 MHz, ($CDCl_3$): 1.24 (t, J = 7.5 Hz, 9H, C(1)H), 2.83 (q, J = 7.5 Hz, 6H, C(2)H), 3.88 (s, 6H, C(5)H2), [lit.[125] (400 MHz,$CDCl_3$) δ 1.23, 2.82 , 3.87]; **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 16.8 (C(1)H2), 22.6 (C(2)H), 39.7 (C(5)H), 137.4 (C3), 140.4 (C4); **LRMS:** (ESI^+) Calculated for $C_{15}H_{27}N_3Na$: 272.2, found $[M+Na]^+$: 272.2.

1,3,5-triethyl-2,4,6-tris(isocyanatomethyl)benzene (151)



Under an inert N_2 atmosphere, a flask was charged with triphosgene (180 mg, 0.61 mmol) and anhydrous toluene (6 mL). A solution of triamine **154** (50 mg, 0.20 mmol) in anhydrous toluene (6 mL) was added dropwise and the reaction mixture was heated to reflux then stirred for 2 hours. The mixture was cooled and the solvent removed under vacuum. The crude solid was triturated with cold toluene, filtered and dried under high vacuum to afford tris-isocyanate **151** (63 mg, 0.20 mmol, 91%) as a colourless oil that crystallised over time.**^{1}H NMR:** (400 MHz, ($CDCl_3$): 1.26 (t, *J* = 7.6 Hz, 9H, C(1)H), 2.84 (q, *J* = 7.6 Hz, 6H, C(2)H), 4.50 (s, 6H, C(5)H_2); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 16.1 (C(1)H_2), 22.8 (C(2)H), 40.4 (C(5)H), 123.1 (C6), 132.4 (C3), 143.1 (C4); **ν_{max}** 2973, 2933, 2875, 2243, 1495, 1453, 1335, 1042, 856, 577 cm^{-1}; **HRMS:** (ESI^+) Calculated for $C_{18}H_{21}N_3O_3Na^+$: 350.1475, found $[M+Na]^+$: 350.1474

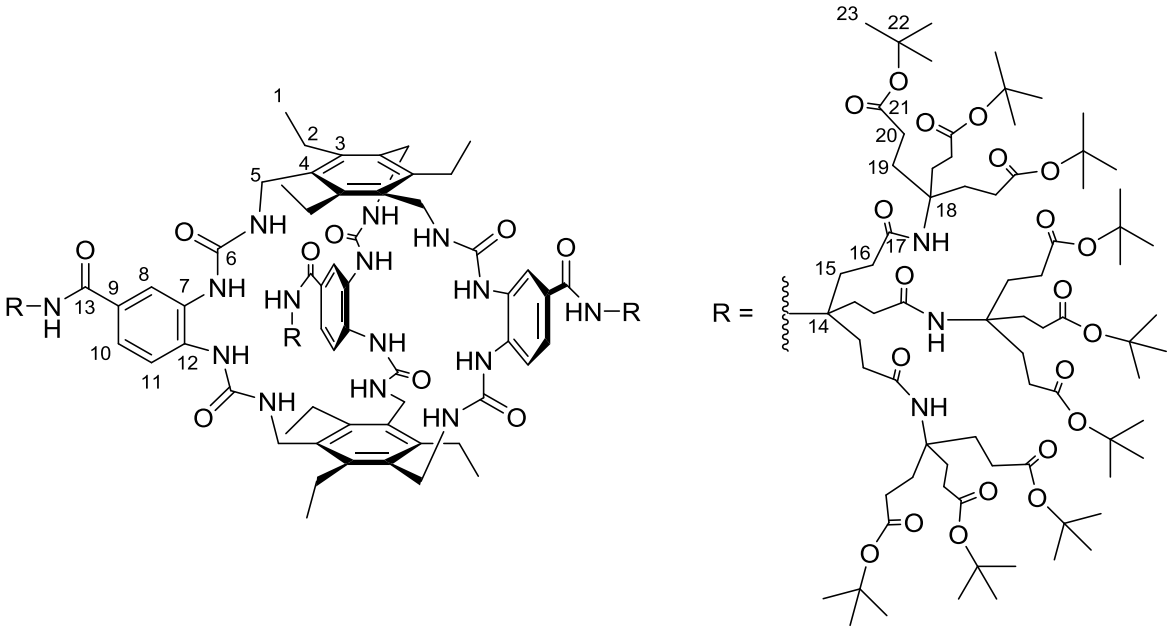
Tris-amino *tert*-butyl G2MM TEB half receptor (152)



Under an inert N_2 atmosphere, **113** (556 mg, 0.31 mmol) was added to a solution of **151** (28 mg, 0.086 mmol) in anhydrous dichloromethane (5 mL). Anhydrous pyridine (40 µL) was added and the reaction heated to 34 °C for 16 hours, after which it was cooled to room temperature and the solvent removed under vacuum. The residue was purified by reverse phase flash chromatography (Biotage SNAP Ultra C18 25 µm) eluting with acetone/water (10:90 → 90:10 over 10 column volumes then 100% acetone over 4 column volumes) to afford **153** (400 mg, 0.072 mmol, 84%) as a white solid. Conversion to **153** was confirmed by limited NMR* and high resolution mass spectrometry (ESI+): m/z calculated for $C_{312}H_{471}N_{21}O_{75}Na_3$: 1928.1155, found $[M+3Na]^{3+}$: 1928.1160; calculated for $C_{312}H_{471}N_{21}O_{75}Na_4$: 1451.8339, found $[M+4Na]^{4+}$: 1451.8320. Under an inert N_2 atmosphere, **153** (300 mg, 0.052 mmol) was dissolved in anhydrous dichloromethane (8 mL) and cooled to 0 °C. DBU (50 µL, 0.30 mmol) was added dropwise and the reaction mixture stirred at 0 °C for 1 hour. The solvent was removed under vacuum and the crude product purified by flash column chromatography (4% MeOH:CH_2Cl_2) to afford **152** as a white solid (238 mg, 0.047 mmol, 91%). **^{1}H NMR:** (400 MHz, (CD_3OD): δ 1.26 (t, *J* = 7.6 Hz, C(1)H$_3$), 1.43 (s, 243H, C(23)H$_3$), 1.94 (m, 54H, C(20)H$_2$), 2.09 (m, 18H, C(15)H$_2$), 2.19 (m, 54H, C(19)H$_2$), 2.24 (m, 18H, C(16)H$_2$), 2.82-2.95 (m, 6H, C(2)H$_2$), 4.52 (s, 6H, C(5)H$_2$), 6.17 (br s, 3H, C(5)H$_2$NH), 7.21 (dd, *J* = 2.1, 8.4 Hz, 3H, C(9)H), 7.30 (d, *J* = 2.1 Hz, 3H, C(11)H), 7.40 (d, *J* = 8.4 Hz, 3H, C(8)H), 7.42 (s, 9H, C(18)NH), 7.93 (s, 3H, OC(13)NH); **^{13}C NMR:** (100 MHz, (CDCl3): δ 17.0 (C1), 22.5 (C2), 28.5 (C23), 30.5 (C19), 30.7 (C20), 32.3 (C15), 32.6 (C16), 37.9 (C5), 58.9 (C18), 59.4 (C14), 81.6 (C22), 117.4 (C11), 118.9 (C9) 124.5 (C8), 130.0 (C10), 132.9 (C3), 133.8 (C7), 141.4 (C12), 145.1 (C4), 157.9 (C6), 170.1 (C13), 174.4 (C21), 175.6 (C17); **HRMS:** (ESI+) calculated for $C_{267}H_{441}N_{21}O_{69}Na_3$: 1706.0472, found $[M+3Na]^{3+}$: 1706.0490. **Retention time:** (C18, acetone/water, 70:30 → 96:4 over 30 mins, then 96:4 isocratic over 10 mins, then 100:0 over 5 mins, 16 mL/min flow rate) 31 mins (analytical), 29 mins (preparative).

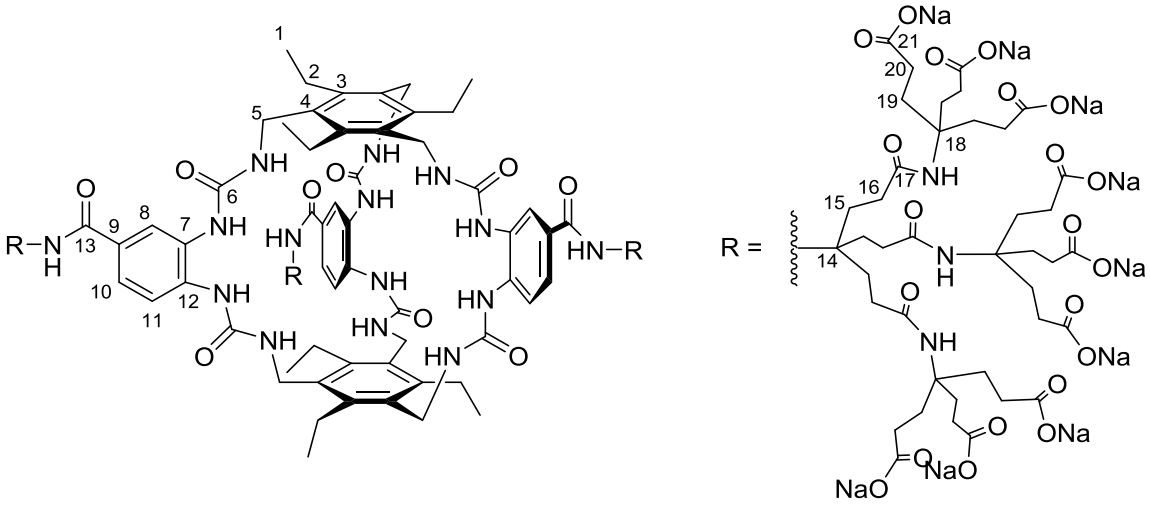
* *Limited NMR studies were only possible due to believed slow conformational exchange of **152** resulting in very broad signals of low intensity.*

***tert* butyl protected G2MM TEB receptor (150)**



Under an inert N_2 atmosphere, **152** (500 mg, 0.1 mmol), octyl β-D-glucoside **12** (59 mg, 0.2 mmol) and DMAP (37 mg, 0.3 mmol) were dissolved in anhydrous dichloromethane (180 mL). A solution of **151** (33 mg, 0.1 mmol) in anhydrous dichloromethane (20 mL) was added at a rate of 1 mL/hr and the reaction stirred at 34 °C for 2 days. The reaction mixture was cooled to room temperature and the solvent removed under vacuum. The residue was purified by preparative HPLC eluting with acetone/water (70:30 → 96:4 over 30 mins, then 96:4 isocratic over 10 mins, then 100:0 over 5 mins). The solvent was removed under vacuum and the product suspended in water and freeze dried to afford **150** as a white solid (257 mg, 0.048 mmol, 48%). **^{1}H NMR:** (400 MHz, (CD_3OD): δ 1.21 (m, 18H C(1)$\underline{H}_3$), 1.43 (s, 243H, C(23)$\underline{H}_3$), 1.95 (m, 54H, C(20)$\underline{H}_2$), 2.13 (m, 18H, C(15)$\underline{H}_2$), 2.20 (m, 54H, C(19)$\underline{H}_2$), 2.25 (m, 18H, C(16)$\underline{H}_2$), 2.72, 2.83 (br m, 6H, C(2)$\underline{H}_2$), 4.40, 4.49 (br s, 6H, C(5)$\underline{H}_2$), 7.43 (s, 9H, N$\underline{H}$), 7.63 (d, *J* = 8.7 Hz, 3H, C(10)$\underline{H}$), 8.04 (d, *J* = 8.7 Hz, 3H, C(11)$\underline{H}$), 8.06 (s, 3H, C(8)$\underline{H}$); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 15.5, 15.6 (C1), 22.3, 22.5 (C2), 27.1 (C23), 29.1 (C19), 29.3 (C20), 30.8 (C15), 31.0 (C16), 37.5 (C5), 57.3 (C18), 58.1 (C14), 80.2 (C22), 120.3 (C11), 123.5 (C8), 124.1 (C10), 127.6 (C7), 129.4 (C9), 132.0, 132.6 (C3), 135.0 (C12), 143.0, 143.2 (C4), 155.8, 156.6 (C6), 168.2 (C13), 173.0 (C21), 174.1 (C17); **HRMS:** (ESI^+) calculated for $C_{285}H_{462}N_{24}O_{72}Na_3$: 1816.1084, found $[M+3Na]^{3+}$: 1816.1093. **Retention time:** (C18, acetone/water, 70:30 → 96:4 over 30 mins, then 96:4 isocratic over 10 mins, then 100:0 over 5 mins, 16 mL/min flow rate) 33 mins (analytical), 31 mins (preparative).

Deprotected G2MM TEB Hexaurea (3)



150 (50 mg, 9.3 µmol) was dissolved in dichloromethane (HPLC grade, 2 ml) and cooled to 0 °C. Trifluoroacetic acid (TFA) (0.5 mL) was added dropwise and the reaction warmed to room temperature and stirred for 16 hours. The solvent was then removed under a flow of nitrogen, then the residue was co-evaporated with toluene (3 x 10 mL) to remove residual TFA, suspended in water and freeze dried. The product was then purified by preparative HPLC eluting with 100% Water (buffered with 0.1 % TFA) → 100% methanol over 40 minutes. The solvent was removed under vacuum, the residue co-evaporated with toluene (3 × 10 mL), the product suspended in water and freeze dried. The solid was then suspended in water, neutralised to pH 7.4 with NaOH (aq), filtered and then freeze dried to afford **3** as a white solid (36.9 mg, 8.3 µmol, 89%). **^{1}H NMR:** (400 MHz, (D_2O): δ 0.93-1.12,1.11-1.25 (br m, 18H, C(1)H_3), 1.89-1.98 (m, 54H, C(20)H_2), 2.08-2.15 (m, 18H, C(15)H_2), 2.13- 2.20 (m, 54H, C(19)H_2), 2.27-2.35 (m, 18H, C(16)H_2), 2.67-2.85 (br m, 12 jH, C(2)H_2), 4.46 (br s, 6H, C(5)H_2), 7.63 (d, *J* = 8.3 Hz, 3H, C(10)H), 7.78 (br s, 3H, C(8)H),7.80-7.85 (br m, 3H, C(11)H); **^{13}C NMR:** (100 MHz, ($CDCl_3$): δ 15.4, 15.5 (C1), 22.4, 22.6 (C2), 30.3 (C19), 30.5 (C15), 30.8 (C16), 31.1 (C20), 37.5, 37.7 (C5), 58.2 (C18), 58.6 (C14),122.4 (C11), 124.9 (C8, C10), 128.4 (C7), 130.0 (C9), 132.0, 132.4 (C3), 135.6 (C12), 144.0, 144.2 (C4), 157.0, 157.7 (C6), 169.6 (C13), 175.2 (C17), 182.1 (C21). **Retention time:** (C18, 100% Water (buffered with 0.1 % TFA) → 100% methanol over 40 minutes, 16 mL/min flow rate) 33 mins (analytical), 32 mins (preparative).

Chapter 8 – References

(1) Varki, A. *Glycobiology* **1993**, *3* (2), 97–130.

(2) Klemm, D.; Heublein, B.; Fink, H.-P.; Bohn, A. *Angew. Chemie Int. Ed.* **2005**, *44* (22), 3358–3393.

(3) Laine, R. A. *Glycobiology* **1994**, *4* (6), 759–767.

(4) Walker, D. B.; Joshi, G.; Davis, A. P. *Cell. Mol. Life Sci.* **2009**, *66* (19), 3177–3191.

(5) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98* (2), 637–674.

(6) Ghazarian, H.; Idoni, B.; Oppenheimer, S. B. *Acta Histochem.* **2011**, *113* (3), 236–247.

(7) Oshovsky, G. V.; Reinhoudt, D. N.; Verboom, W. *Angew. Chemie Int. Ed.* **2007**, *46* (14), 2366–2393.

(8) Davis, A. P.; Kubik, S.; Dalla Cort, A. *Org. Biomol. Chem.* **2015**, *13* (9), 2499–2500.

(9) Davis, A. P.; Wareham, R. S. *Angew. Chemie Int. Ed.* **1999**, *38* (20), 2978–2996.

(10) Davis, A. P.; Wareham, R. S. *Angew. Chemie.* **1999**, *111* (20), 3160–3179.

(11) Toone, E. J. *Curr. Opin. Struct. Biol.* **1994**, *4* (5), 719–728.

(12) Goldstein, I. J.; Hayes, C. E. *Adv. Carbohydr. Chem. Biochem.* **1978**, *35*, 127–340.

(13) Rozwarski, D. A.; Swami, B. M.; Brewer, C. F.; Sacchettini, J. C. *J. Biol. Chem.* **1998**, *273* (49), 32818–32825.

(14) Davis, A. P. *Nature* **2010**, *464* (7286), 169–170.

(15) Sun, X.; James, T. D. *Chem. Rev.* **2015**, *115* (15), 8001–8037.

(16) Thordarson, P. *Chem. Soc. Rev.* **2011**, *40* (3), 1305–1323.

(17) Kleckner, I. R.; Foster, M. P. *Biochim. Biophys. Acta* **2011**, *1814* (8), 942–968.

(18) Klein, E.; Ferrand, Y.; Barwell, N. P.; Davis, A. P. *Angew. Chemie Int. Ed.* **2008**, *47* (14), 2693–2696.

(19) Davis, A. P. *Org. Biomol. Chem.* **2009**, *7* (18), 3629.

(20) Kikuchi, Y.; Tanaka, Y.; Sutarto, S.; Kobayashi, K.; Toi, H.; Aoyama, Y. *J. Am. Chem. Soc.* **1992**, *114* (26), 10302–10306.

(21) Bhattarai, K. M.; Davis, A. P.; Perry, J. .; Walter, C. J.; Menzer, S.; Williams, D. J. *J. Org. Chem.* **1997**, *62* (24), 8463–8473.

(22) Nativi, C.; Cacciarini, M.; Francesconi, O.; Vacca, A.; Moneti, G.; Ienco, A.; Roelens, S. *J. Am. Chem. Soc.* **2007**, *129* (14), 4377–4385.

(23) Francesconi, O.; Gentili, M.; Roelens, S. *J. Org. Chem.* **2012**, *77* (17), 7548–7554.

(24) Mazik, M.; Cavga, H.; Jones‡, P. G. *J. Am. Chem. Soc.* **2005**, *127* (25), 9045–9052.

(25) Lippe, J.; Mazik, M. *J. Org. Chem.* **2015**, *80* (3), 1427–1439.

(26) Chandramouli, N.; Ferrand, Y.; Lautrette, G.; Kauffmann, B.; Mackereth, C. D.; Laguerre, M.; Dubreuil, D.; Huc, I. *Nat. Chem.* **2015**, *7* (4), 334–341.

(27) Kataev, E. A.; Müller, C. *Tetrahedron* **2014**, *70* (2), 137–167.

(28) Yoon, J.; Czarnik, A. W. *J. Am. Chem. Soc.* **1992**, *114* (14), 5874–5875.

(29) James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S. *Angew. Chemie Int. Ed. English* **1996**, *35* (17), 1910–1922.

(30) Lorand, J. P.; Edwards, J. O. *J. Org. Chem.* **1959**, *24* (6), 769–774.

(31) Wu, X.; Li, Z.; Chen, X.-X.; Fossey, J. S.; James, T. D.; Jiang, Y.-B. *Chem. Soc. Rev.* **2013**, *42* (20), 8032.

(32) Cambre, J. N.; Sumerlin, B. S. *Polymer (Guildf).* **2011**, *52* (21), 4631–4643.

(33) Ferrand, Y.; Klein, E.; Barwell, N. P.; Crump, M. P.; Jiménez-Barbero, J.; Vicent, C.; Boons, G.-J.; Ingale, S.; Davis, A. P. *Angew. Chemie Int. Ed.* **2009**, *48* (10), 1775–1779.

(34) Barwell, N. P.; Crump, M. P.; Davis, A. P. *Angew. Chemie Int. Ed.* **2009**, *48* (41), 7673–7676.

(35) Ke, C.; Destecroix, H.; Crump, M. P.; Davis, A. P. *Nat. Chem.* **2012**, *4* (9), 718–723.

(36) Destecroix, H.; Renney, C. M.; Mooibroek, T. J.; Carter, T. S.; Stewart, P. F. N.; Crump, M. P.; Davis, A. P. *Angew. Chemie Int. Ed.* **2015**, *54* (7), 2057–2061.

(37) Sookcharoenpinyo, B.; Klein, E.; Ferrand, Y.; Walker, D. B.; Brotherhood, P. R.; Ke, C.; Crump, M. P.; Davis, A. P. *Angew. Chemie Int. Ed.* **2012**, *51* (19), 4586–4590.

(38) Ferrand, Y.; Crump, M. P.; Davis, A. P. *Science* **2007**, *318* (5850), 619–622.

(39) Mooibroek, T. J.; Casas-Solvas, J. M.; Harniman, R. L.; Renney, C. M.; Carter, T. S.; Crump, M. P.; Davis, A. P. *Nat. Chem.* **2016**, *8* (1), 69–74.

(40) Ríos, P.; Mooibroek, T. J.; Carter, T. S.; Williams, C.; Wilson, M. R.; Crump, M. P.; Davis, A. P. *Chem. Sci.* **2017**, *8* (5), 4056–4061.

(41) Tang, W. J.; Fernandez, J.; Sohn, J. J.; Amemiya, C. T. *Curr. Biol.* **2015**, *25* (7), 897–900.

(42) Gao, D.; Chundawat, S. P. S.; Sethi, A.; Balan, V.; Gnanakaran, S.; Dale, B. E. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (27), 10922–10927.

(43) Chundawat, S. P. S.; Beckham, G. T.; Himmel, M. E.; Dale, B. E. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2* (1), 121–145.

(44) Chen, J.-K.; Shen, C.-R.; Liu, C.-L. *Mar. Drugs* **2010**, *8* (9), 2493–2516.

(45) Huang, Y.-B.; Fu, Y. *Green Chem.* **2013**, *15* (5), 1095.

(46) Renney, C. M. *PhD Thesis*, University of Bristol, **2014**.

(47) Schnabel, W. J.; Kober, E. *J. Org. Chem.* **1969**, *34* (4), 1162–1165.

(48) Dias, C. M. *PhD Thesis*, University of Bristol, **2018**.

(49) Gassensmith, J. J.; Arunkumar, E.; Barr, L.; Baumes, J. M.; DiVittorio, K. M.; Johnson, J. R.; Noll, B. C.; Smith, B. D. *J. Am. Chem. Soc.* **2007**, *129* (48), 15054–15059.

(50) Moodie, R. B.; Sansom, P. J. *J. Chem. Soc. Perkin Trans. 2* **1981**, No. 4, 664.

(51) Wasserman, H. H.; Scheffer, J. R.; Cooper, J. L. *J. Am. Chem. Soc.* **1972**, *94* (14), 4991–4996.

(52) Bregnhøj, M.; Westberg, M.; Jensen, F.; Ogilby, P. R. *Phys. Chem. Chem. Phys.* **2016**, *18* (33), 22946–22961.

(53) Salokhiddinov, K. I.; Byteva, I. M.; Gurinovich, G. P. *J. Appl. Spectrosc.* **1981**, *34* (5), 561–564.

(54) Lundt, B. F.; Johansen, N. L.; Vølund, A.; Markussen, J. *Int. J. Pept. Protein Res.* **2009**, *12* (5), 258–268.

(55) Wolff, H. *Org. React.* **1946**, *3*, 307.

(56) Sandin, R. B.; Melby, R.; Crawford, R.; McGreer, D. *J. Am. Chem. Soc.* **1956**, *78* (15), 3817–3819.

(57) Filali, E.; Lloyd-Jones, G.; Sale, D. *Synlett* **2009**, *2009* (2), 205–208.

(58) Brooks, C. A. G.; Davis, K. M. C.; Blandamer, M. J. *J. Solution Chem.* **1974**, *3* (3), 247–250.

(59) Chen, B.; Sun, H.-X.; Qin, J.-F.; Wang, B. *Tetrahedron Lett.* **2016**, *57* (3), 253–255.

(60) Rios, P.; Carter, T. S.; Mooibroek, T. J.; Crump, M. P.; Lisbjerg, M.; Pittelkow, M.; Supekar, N. T.; Boons, G.-J.; Davis, A. P. *Angew. Chemie Int. Ed.* **2016**, *55* (10), 3387–3392.

(61) Kojima, H.; Nakatsubo, N.; Kikuchi, K.; Kawahara, S.; Kirino, Y.; Nagoshi, H.; Hirata, Y.; Nagano, T. *Anal. Chem.* **1998**, *70* (13), 2446–2453.

(62) Newkome, G. R.; Weis, C. D. *Org. Prep. Proced. Int.* **1996**, *28* (4), 495–498.

(63) Reddy, P. Y.; Kondo, S.; Toru, T.; Ueno, Y. *J. Org. Chem.* **1997**, *62* (8), 2652–2654.

(64) Polenz, I.; Laue, A.; Uhrin, T.; Rüffer, T.; Lang, H.; Schmidt, F. G.; Spange, S. *Polym. Chem.* **2014**, *5* (23), 6678–6686.

(65) Cates, L. A.; Li, V.-S.; Basrur, J. P.; Alkadhi, K. A. *J. Pharm. Sci.* **1986**, *75* (4), 407–409.

(66) Galy, J.-P.; Vidal, R.; Pierre Galy, J.; Jean Vincent, E.; Marie Galy, A.; Barbe, J. *Heterocycles* **1986**, *24* (5), 1419.

(67) Blanco-Canosa, J. B.; Dawson, P. E. *Angew. Chemie* **2008**, *120* (36), 6957–6961.

(68) Sugimoto, A.; Yamano, J.; Suyama, K.; Yoneda, S. *J. Chem. Soc. Perkin Trans. 1* **1989**, *0* (3), 483.

(69) Audu, A. A.; Heyn, A. H. A. *Water Res.* **1988**, *22* (9), 1155–1162.

(70) Dam, T. K.; Brewer, C. F. In *Lectins*; Elsevier, 2007; pp 75–101.

(71) Ahmad, M.; Kalinina, O.; Lengauer, T. *J. Cheminform.* **2014**, *6* (Suppl 1), P35.

(72) Grünberg, R.; Nilges, M.; Leckner, J. *Structure* **2006**, *14* (4), 683–693.

(73) Tedesco, D.; Bertucci, C. *J. Pharm. Biomed. Anal.* **2015**, *113*, 34–42.

(74) McIntosh, M.; Stone, B. A.; Stanisich, V. A. *Appl. Microbiol. Biotechnol.* **2005**, *68* (2), 163–173.

(75) Manners, D. J.; Masson, A. J.; Patterson, J. C. *Biochem. J.* **1973**, *135* (1), 19–30.

(76) Odabasi, Z.; Mattiuzzi, G.; Estey, E.; Kantarjian, H.; Saeki, F.; Ridge, R. J.; Ketchum, P. A.; Finkelman, M. A.; Rex, J. H.; Ostrosky-Zeichner, L. *Clin. Infect. Dis.* **2004**, *39* (2), 199–205.

(77) Ostrosky-Zeichner, L.; Alexander, B. D.; Kett, D. H.; Vazquez, J.; Pappas, P. G.; Saeki, F.; Ketchum, P. A.; Wingard, J.; Schiff, R.; Tamura, H.; Finkelman, M. A.; Rex, J. H. *Clin. Infect. Dis.* **2005**, *41* (5), 654–659.

(78) Arnaud, J.; Audfray, A.; Imberty, A. *Chem. Soc. Rev.* **2013**, *42* (11), 4798.

(79) Mergenthaler, P.; Lindauer, U.; Dienel, G. A.; Meisel, A. *Trends Neurosci.* **2013**, *36* (10), 587–597.

(80) Aleshin, A. E.; Zeng, C.; Bourenkov, G. P.; Bartunik, H. D.; Fromm, H. J.; Honzatko, R. B. *Structure* **1998**, *6* (1), 39–50.

(81) Solís, D.; Bovin, N. V.; Davis, A. P.; Jiménez-Barbero, J.; Romero, A.; Roy, R.; Smetana, K.; Gabius, H.-J. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850* (1), 186–235.

(82) Clark, N. G.; Fox, K. M.; Grandy, S.; SHIELD Study Group. *Diabetes Care* **2007**, *30* (11), 2868–2873.

(83) Wu, Q.; Wang, L.; Yu, H.; Wang, J.; Chen, Z. *Chem. Rev.* **2011**, *111* (12), 7855–7875.

(84) Bakh, N. A.; Cortinas, A. B.; Weiss, M. A.; Langer, R. S.; Anderson, D. G.; Gu, Z.; Dutta, S.; Strano, M. S. *Nat. Chem.* **2017**, *9* (10), 937–943.

(85) Cobelli, C.; Renard, E.; Kovatchev, B. *Diabetes* **2011**, *60* (11), 2672–2682.

(86) Brownlee, M.; Cerami, A. *Diabetes* **1983**, *32* (6), 499–504.

(87) Wang, H.-C.; Lee, A.-R. *J. Food Drug Anal.* **2015**, *23* (2), 191–200.

(88) Shoham, J.; Inbar, M.; Sachs, L. *Nature* **1970**, *227* (5264), 1244–1246.

(89) Harris, J. M.; Reyes, C.; Lopez, G. P. *J. Diabetes Sci. Technol.* **2013**, *7* (4), 1030–1038.

(90) Bisson, A. P.; Lynch, V. M.; Monahan, M.-K. C.; Anslyn, E. V. *Angew. Chemie Int. Ed. English* **1997**, *36* (21), 2340–2342.

(91) Snowden, T. S.; Bisson, A. P.; Anslyn, E. V. *J. Am. Chem. Soc.* **1999**, *121* (26), 6324–6325.

(92) Wang, X.; Hof, F. *Beilstein J. Org. Chem.* **2012**, *8*, 1–10.

(93) Joshi, G. *PhD Thesis*, University of Bristol, **2012**.

(94) Bhattar, S. L.; Kolekar, G. B.; Patil, S. R. *J. Dispers. Sci. Technol.* **2010**, *32* (1), 23–27.

(95) Charalambides, Y. C.; Moratti, S. C. *Synth. Commun.* **2007**, *37* (6), 1037–1044.

(96) Renney, C. M.; Fukuhara, G.; Inoue, Y.; Davis, A. P. *Chem. Commun.* **2015**, *51* (46), 9551–9554.

(97) Hugo, W. B. *J. Appl. Bacteriol.* **1991**, *71* (1), 9–18.

(98) Asensio, J. L.; Ardá, A.; Cañada, F. J.; Jiménez-Barbero, J. *Acc. Chem. Res.* **2013**, *46* (4), 946–954.

(99) Geraldes, C. F. G. C.; Laurent, S. *Contrast Media Mol. Imaging* **2009**, *4* (1), 1–23.

(100) Turnbull, W. B.; Daranas, A. H. *J. Am. Chem. Soc.* **2003**, *125* (48), 14859–14866.

(101) Sookcharoenpinyo, B.; Klein, E.; Ke, C.; Davis, A. P. *Supramol. Chem.* **2013**, *25* (9–11), 650–655.

(102) Basu, A.; Slama, M. Q.; Nicholson, W. T.; Langman, L.; Peyser, T.; Carter, R.; Basu, R. *J. Diabetes Sci. Technol.* **2017**, *11* (5), 936–941.

(103) Larner, J. *Int. J. Exp. Diabetes Res.* **2002**, *3* (1), 47–60.

(104) Gerasimenko, J. V; Flowerdew, S. E.; Voronina, S. G.; Sukhomlin, T. K.; Tepikin, A. V; Petersen, O. H.; Gerasimenko, O. V. *J. Biol. Chem.* **2006**, *281* (52), 40154–40163.

(105) Ma, K.; Thomason, L. A. M.; McLaurin, J. *Adv. Pharmacol.* **2012**, *64*, 177–212.

(106) Loewus, F. A.; Loewus, M. W. *Annu. Rev. Plant Physiol.* **1983**, *34* (1), 137–161.

(107) Pombinho, A. R.; Laizé, V.; Molha, D. M.; Marques, S. M. P.; Cancela, M. L. *Cell Tissue Res.* **2004**, *315* (3), 393–406.

(108) Mitaka, T.; Sattler, G. L.; Pitot, H. C. *J. Cell. Physiol.* **1991**, *147* (3), 495–504.

(109) Tajmir-Riahi, H.-A. *Carbohydr. Res.* **1988**, *183* (1), 35–46.

(110) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2004**, *108* (14), 4546–4557.

(111) Krebs, H. A. *Annu. Rev. Biochem.* **1950**, *19* (1), 409–430.

(112) Burmeister, J. J.; Arnold, M. A. *Anal. Lett.* **1995**, *28* (4), 581–592.

(113) Linek, V.; Beneš, P.; Sinkule, J.; Holeček, O.; Malý, V. *Biotechnol. Bioeng.* **1980**, *22* (12), 2515–2527.

(114) Winiewska, M.; Bugajska, E.; Poznański, J. *PLoS One* **2017**, *12* (3), e0173260.

(115) Hunter, C. A. *Angew. Chemie Int. Ed.* **2004**, *43* (40), 5310–5324.

(116) Horn, J.; Schanda, J.; Friess, W. *Eur. J. Pharm. Biopharm.* **2018**.

(117) Allenmark, S. *Chirality* **2003**, *15* (5), 409–422.

(118) Wang, R.; Zhang, M.; Liu, X.; Zhang, L.; Kang, Z.; Wang, W.; Wang, X.; Dai, F.; Sun, D. *Inorg. Chem.* **2015**, *54* (13), 6084–6086.

(119) Chung, Y.; Duerr, B. F.; McKelvey, T. A.; Nanjappan, P.; Czarnik, A. W. *J. Org. Chem.* **1989**, *54* (5), 1018–1032.

(120) Kim, S. K.; Singh, N. J.; Kim, S. J.; Swamy, K. M. K.; Kim, S. H.; Lee, K.-H.; Kim, K. S.; Yoon, J. *Tetrahedron* **2005**, *61* (19), 4545–4550.

(121) Turoczi, M.-C.; Simon, M.; Badea, V.; Csunderlik, C. *Molecules* **2008**, *13* (12), 3192–3197.

(122) Chattopadhyay, P.; Nagpal, R.; Pandey, P. S. *Aust. J. Chem.* **2008**, *61* (3), 216.

(123) Momose, T.; Inaba, A.; Inoue, K.; Miyahara, K.; Mori, T. *Chem. Pharm. Bull. (Tokyo).* **1964**, *12* (1), 14–18.

(124) Ruan, Y.; Wang, B.-Y.; Erb, J. M.; Chen, S.; Hadad, C. M.; Badjić, J. D. *Org. Biomol. Chem.* **2013**, *11* (44), 7667.

(125) Moragues, M. E.; Santos-Figueroa, L. E.; Ábalos, T.; Sancenón, F.; Martínez-Máñez, R. *Tetrahedron Lett.* **2012**, *53* (38), 5110–5113.

Chapter9 – Appendix

9.1 Anthracene Receptor (1)

9.1.1 Anthracene receptor (1) studies

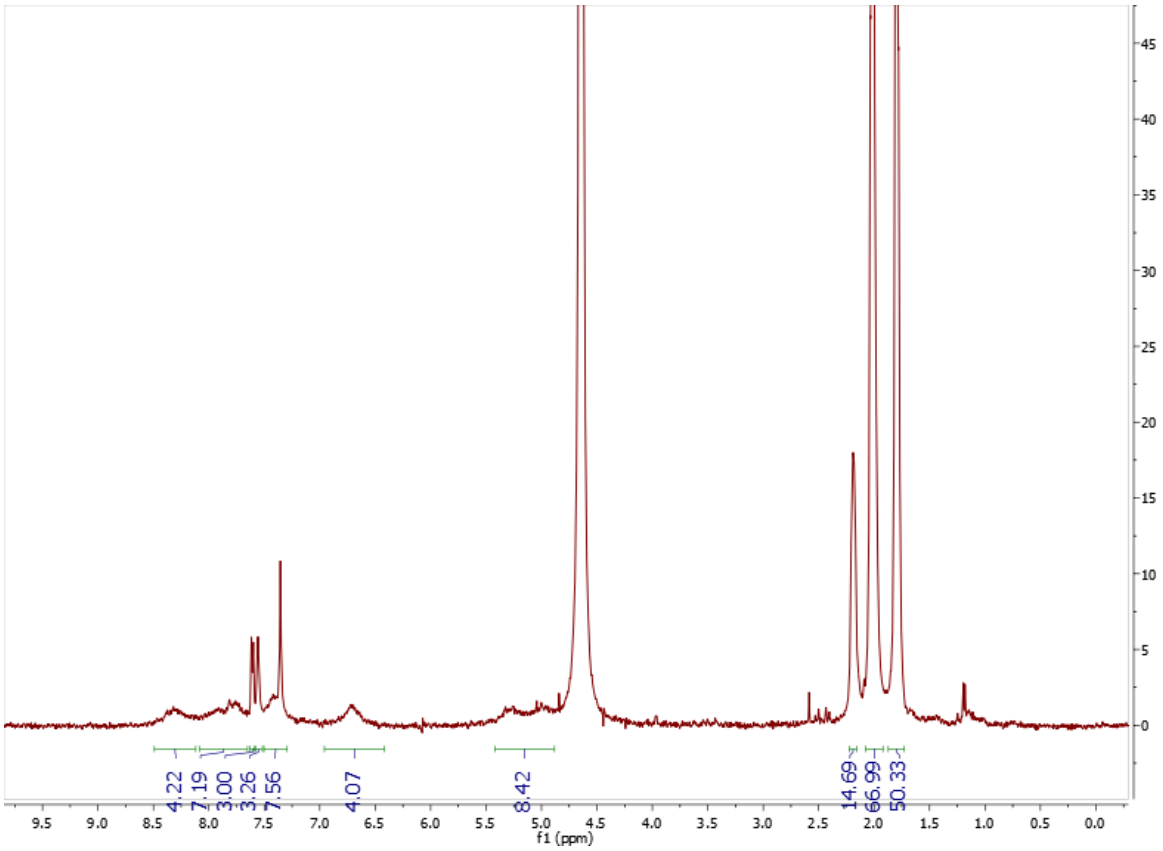


Figure 105 ^{1}H NMR spectrum of receptor **1** in D_2O (pH 7.4) at 298 K.

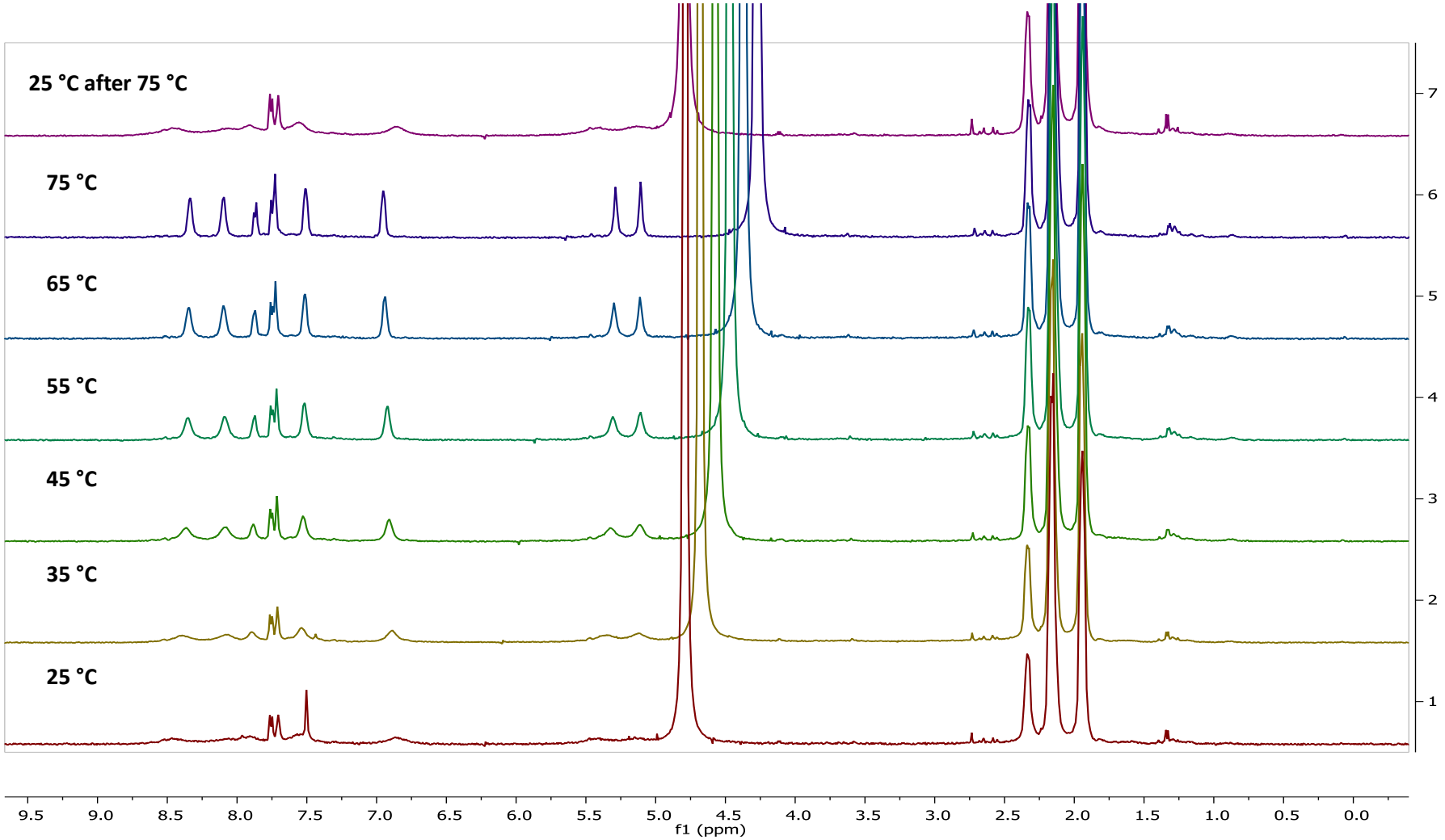


Figure 106 Partial ^{1}H NMR spectra showing receptor **1** at various temperatures in D_2O (pH 7.4). Receptor signals sharpen upon heating, most notably for anthracene and urea methylene protons.

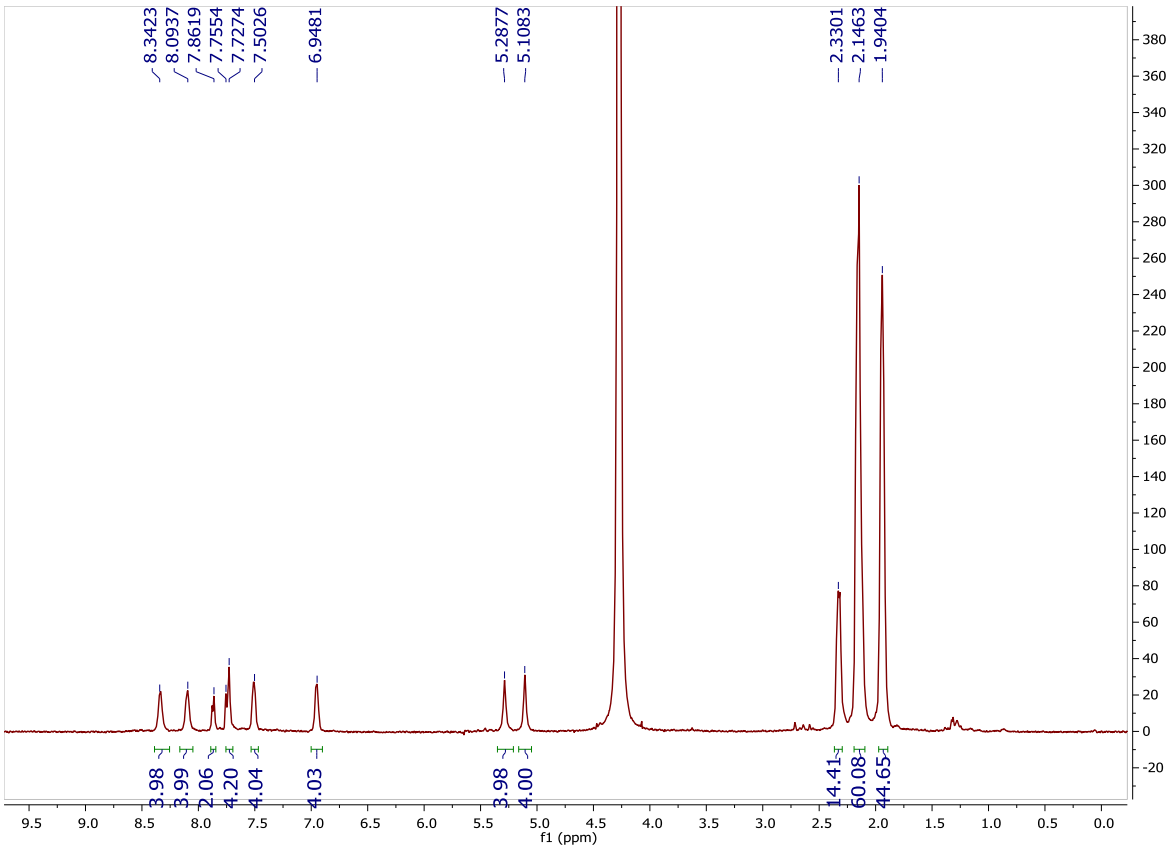


Figure 107 ^{1}H NMR spectrum of receptor **1** in D_2O (pH 7.4) at 348 K.

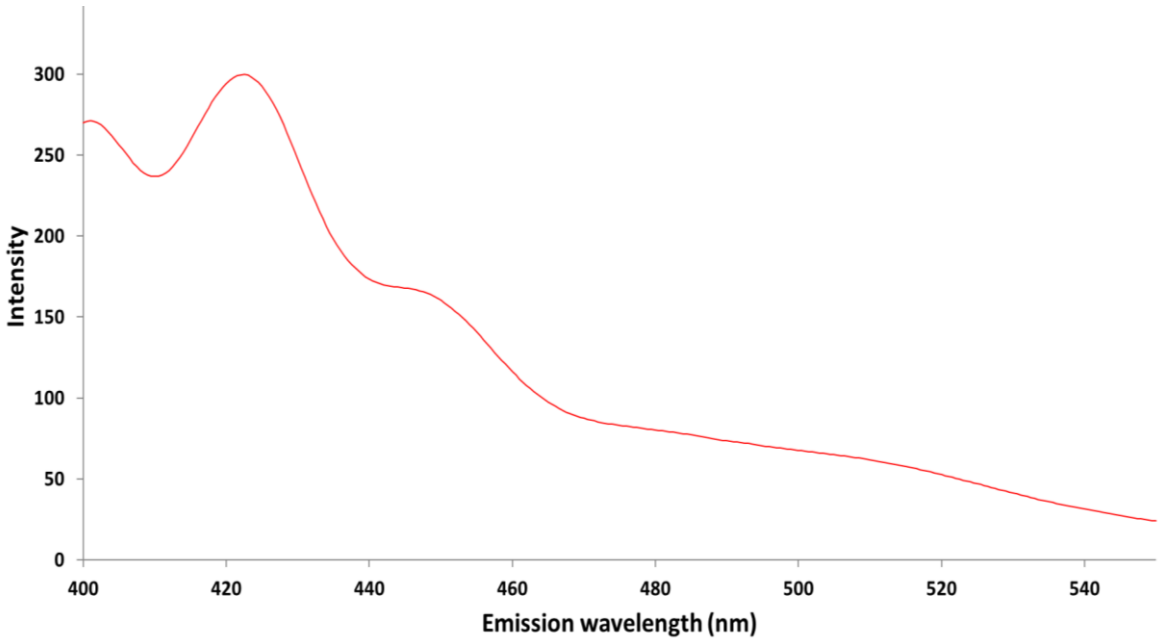


Figure 108 Fluorescence emission spectrum of receptor **1** (10 µM) in water (pH 7.4) at 298 K, using excitation wavelength of 395 nm.

9.1.2 Binding studies for anthracene receptor (1)

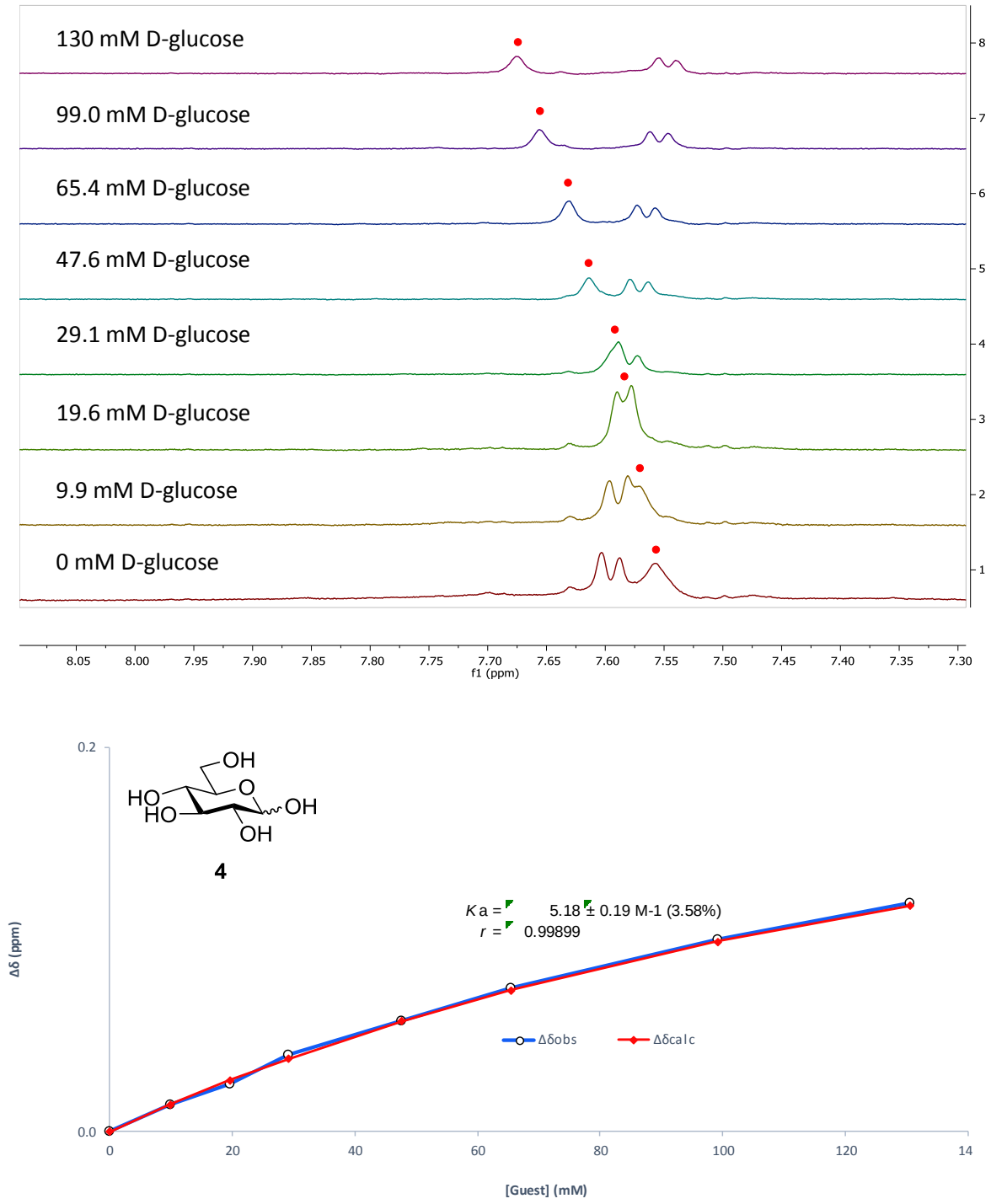


Figure 109 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for **1** (1 mM) titrated with a combined solution of D-glucose **4** (1 M) and **1** (1 mM), in D_2O with at pH 7.4 and 298 K. Change in chemical shifts (Δδ, ppm) denoted with ● were plotted against D-glucose **4** concentration (mM). The calculated values for Δδ are overlaid with the observed values, giving K_a = 5.1 ± 0.2 M^{-1} (3.6%).

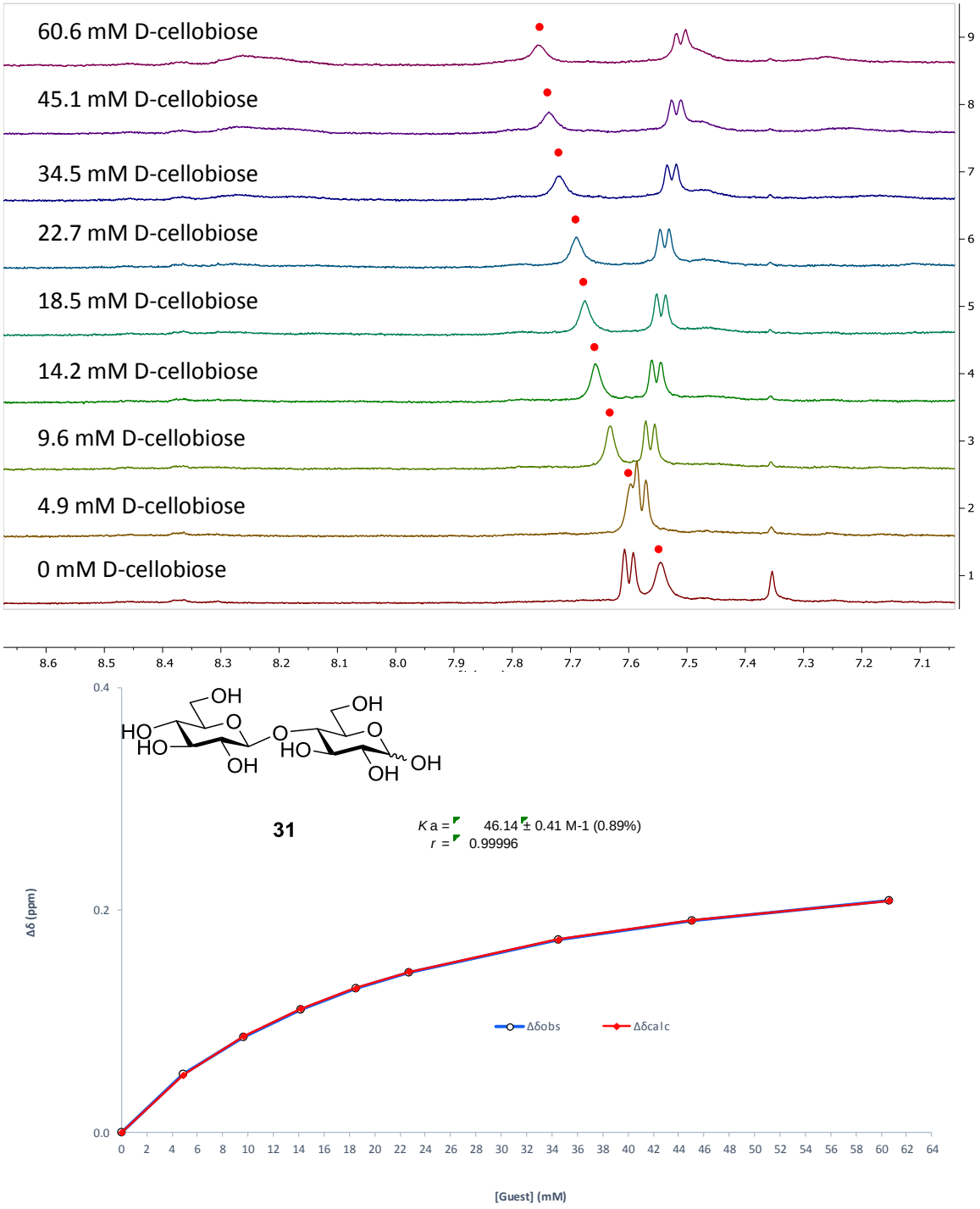


Figure 110 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for **1** (0.25 mM) titrated with a combined solution of D-cellobiose **31** (250 mM) and **1** (0.25 mM), in D_2O with at pH 7.4 and 298 K. Change in chemical shifts (Δδ, ppm) denoted with ● were plotted against D-cellobiose concentration (mM). The calculated values for Δδ are overlaid with the observed values, giving K_a = 46 ± 0.4 M^{-1} (0.89%).

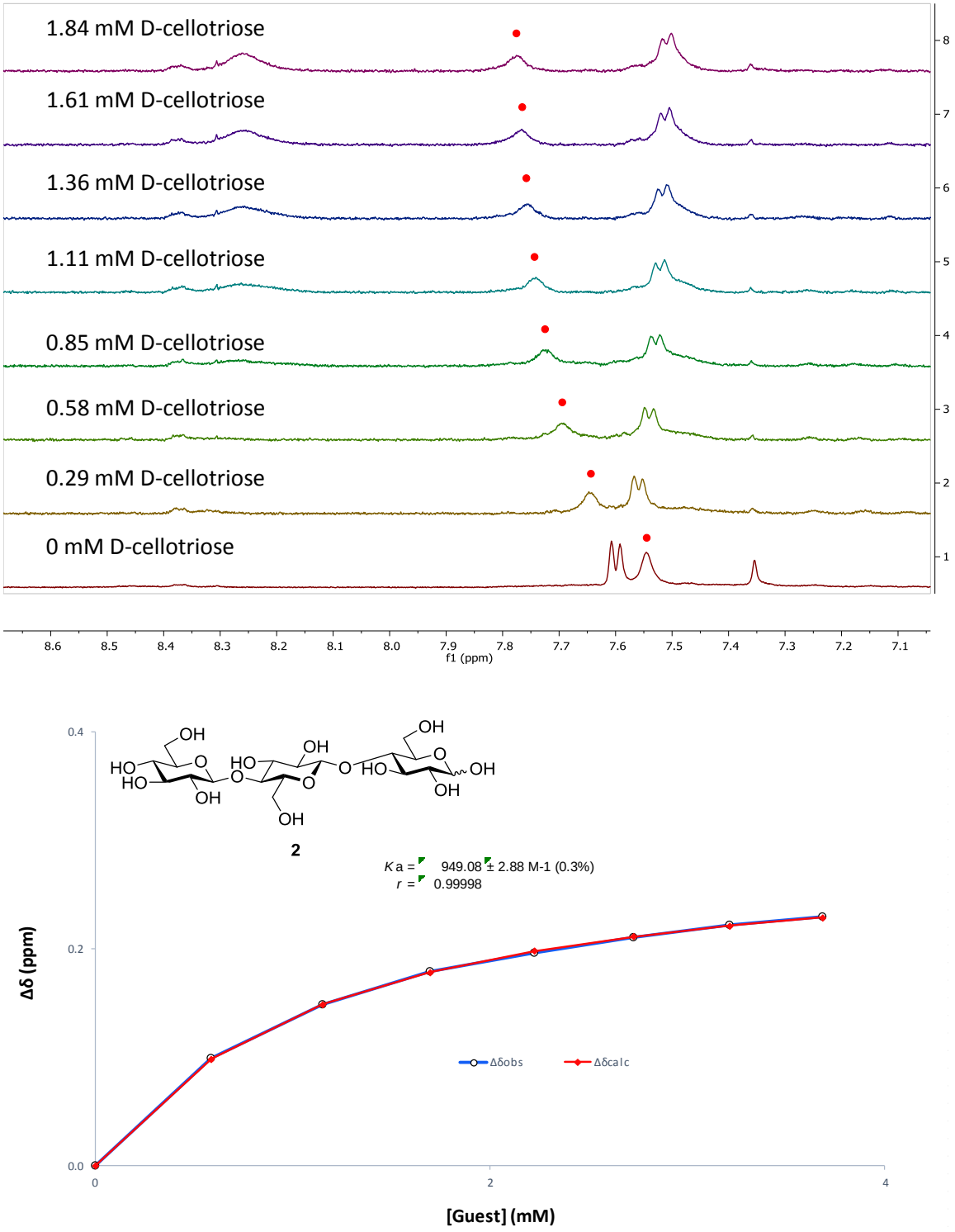
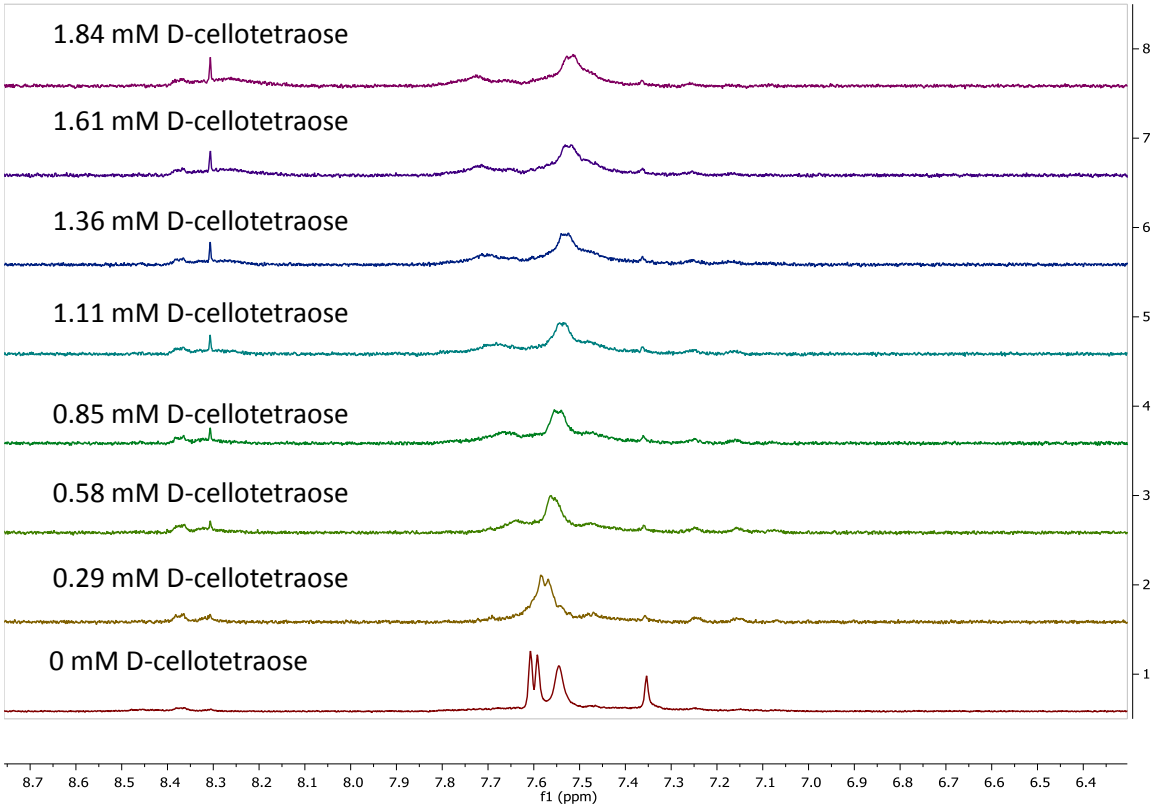
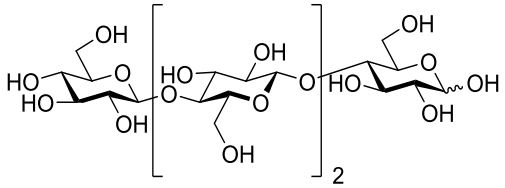


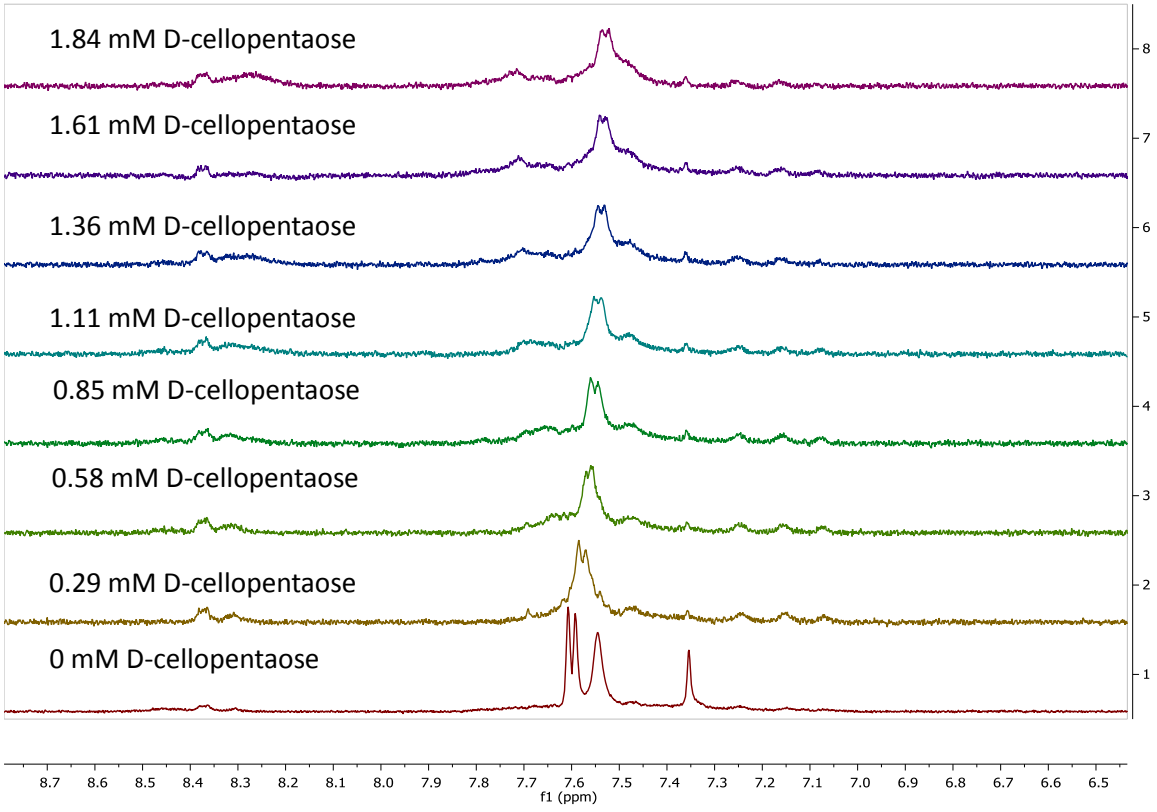
Figure 111 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for **1** (0.2 mM) titrated with a combined solution of D-cellotriose **2** (15 mM) and **1** (0.2 mM), in D_2O with at pH 7.4 and 298 K. Change in chemical shifts (Δδ, ppm) denoted with ● were plotted against D-cellotriose **2** concentration (mM). The calculated values for Δδ are overlaid with the observed values, giving K_a = 949 ± 2.9 M-1 (0.3%).

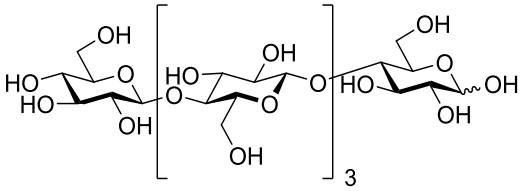




33

Figure 112 Partial $_1$H NMR spectra for **1** (0.2 mM) titrated with a combined solution of D-cellotetraose **33** (15 mM) and **1** (0.2 mM), in D_2O with at pH 7.4 and 298 K. Spectra imply binding with intermediate rate of exchange, thus no K_a was determinable.





130

Figure 113 Partial ^{1}H NMR spectra for **1** (0.2 mM) titrated with a combined solution of D-cellopentaose **130** (15 mM) and **1** (0.2 mM), in D_2O with at pH 7.4 and 298 K. Spectra imply binding with intermediate rate of exchange, thus no K_a was determinable.

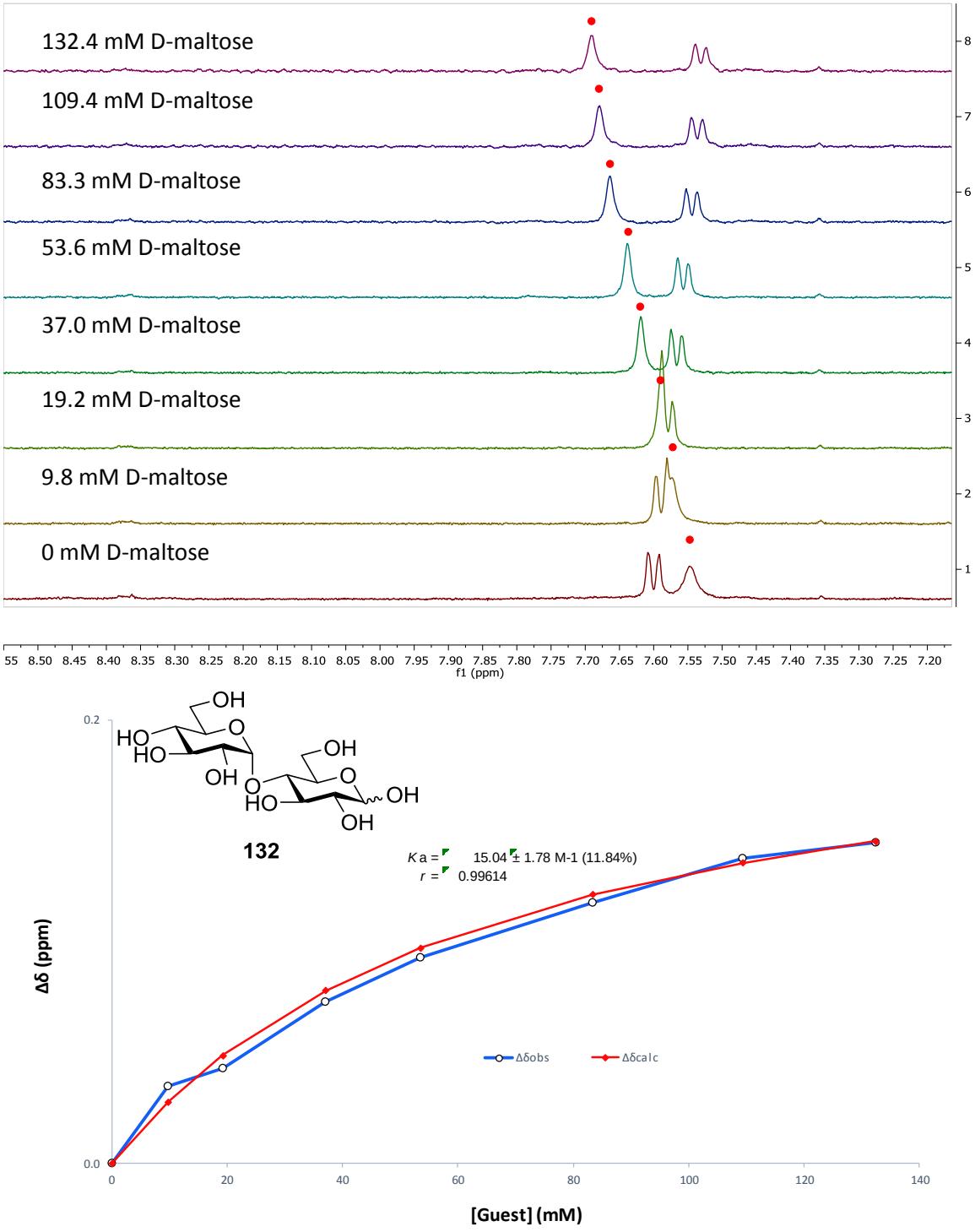


Figure 114 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for **1** (0.2 mM) titrated with a combined solution of D-maltose **132** (500 mM) and 90 (0.2 mM), in D_2O with at pH 7.4 and 298 K. Change in chemical shifts (Δδ, ppm) denoted with ● were plotted against D-maltose concentration (mM). The calculated values for Δδ are overlaid with the observed values, giving K_a = 15 ± 1.8 M^{-1} (11.8%).

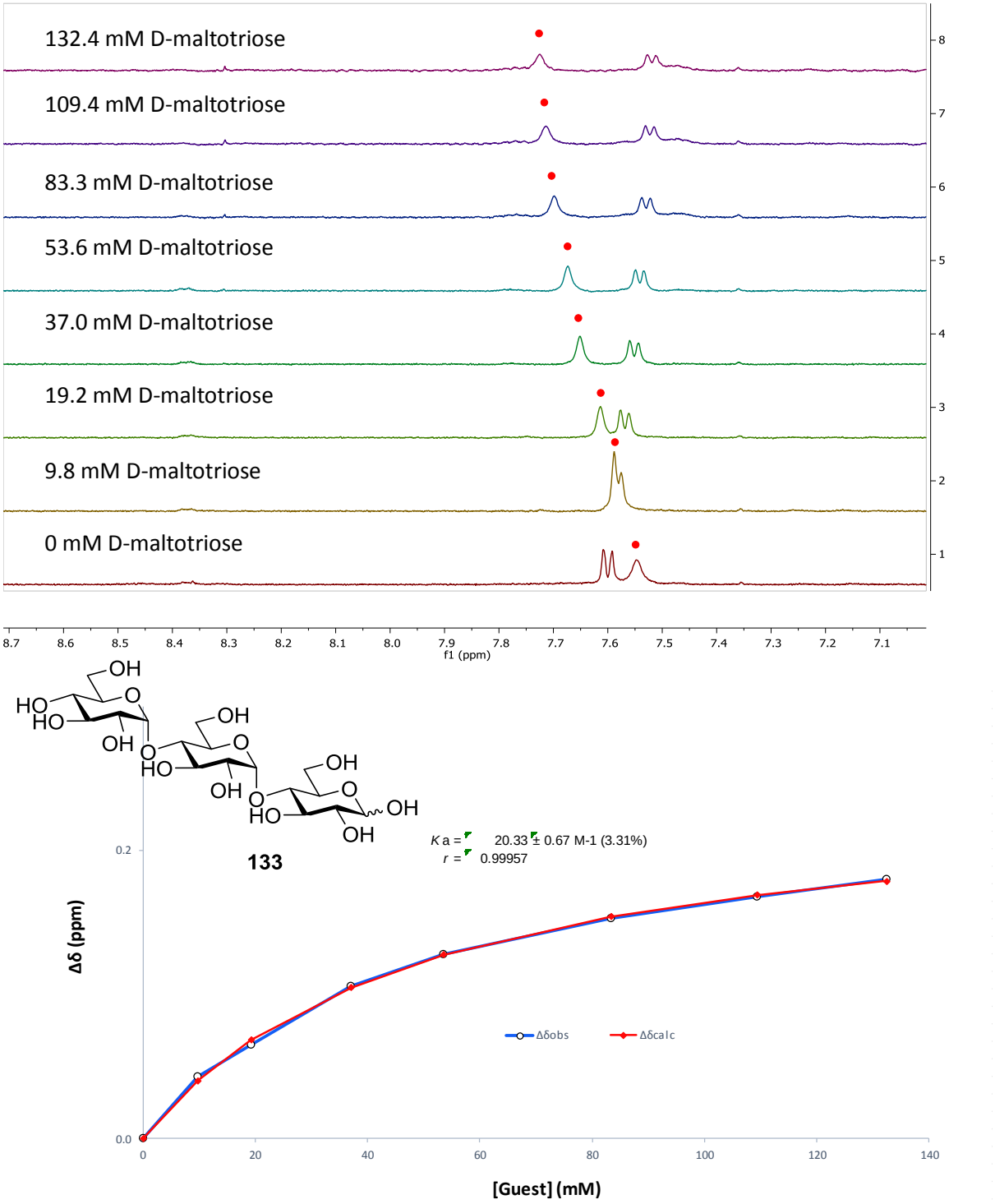


Figure 115 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for **1** (0.2 mM) titrated with a combined solution of D-maltotriose **133** (500 mM) and **1** (0.2 mM), in D_2O with at pH 7.4 and 298 K. Change in chemical shifts (Δδ, ppm) denoted with ● were plotted against D-maltotriose concentration (mM). The calculated values for Δδ are overlaid with the observed values, giving K_a = 20 ± 0.7 M^{-1} (3.3%).

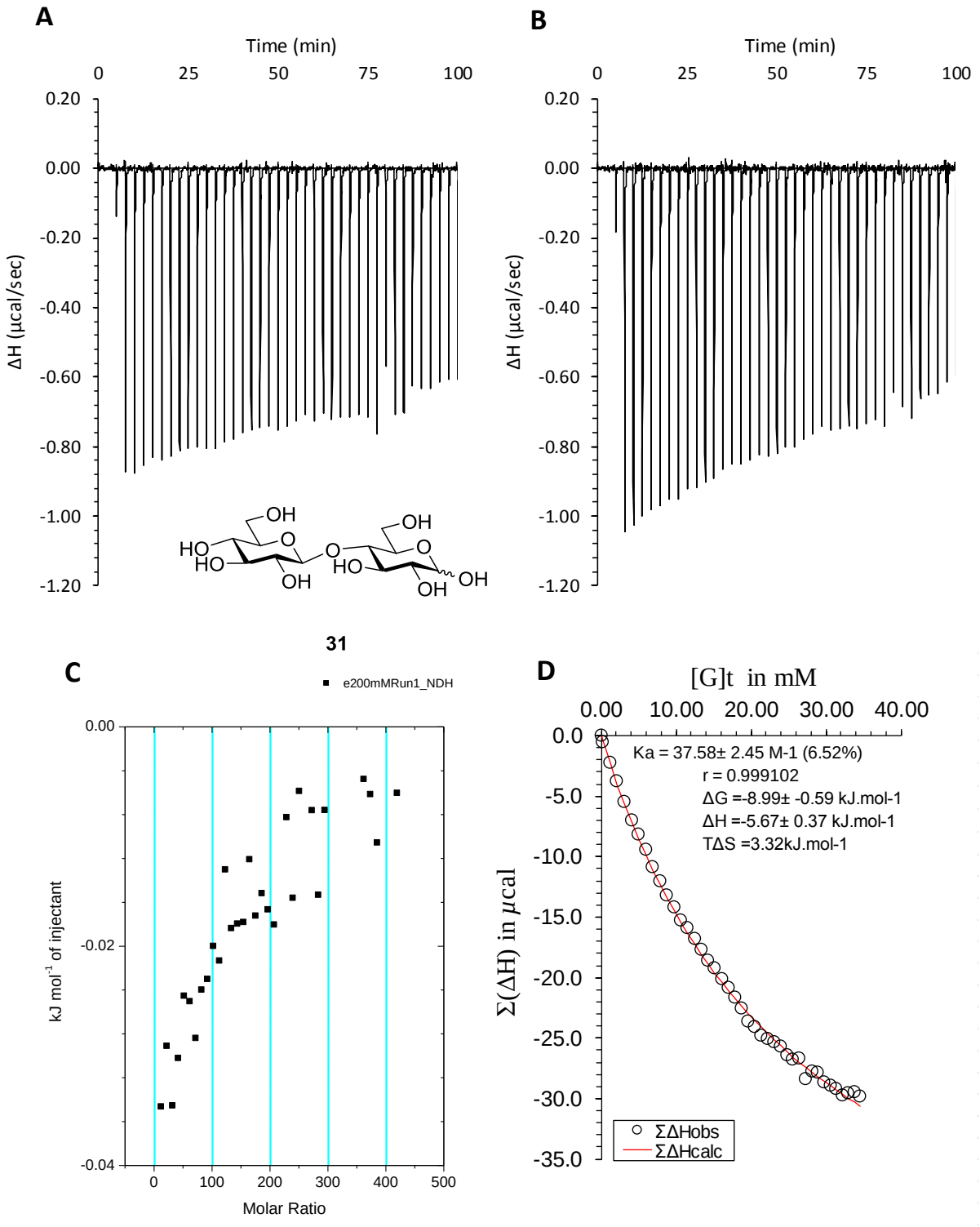


Figure 116 ITC binding results for **1** (0.2 mM) titrated with D-cellobiose (200 mM) in water at 298 K, in which: A) shows the blank run (addition of substrate into water); B) shows the titration (substrate into receptor); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 37.6 ± 2.5 M^{-1}).

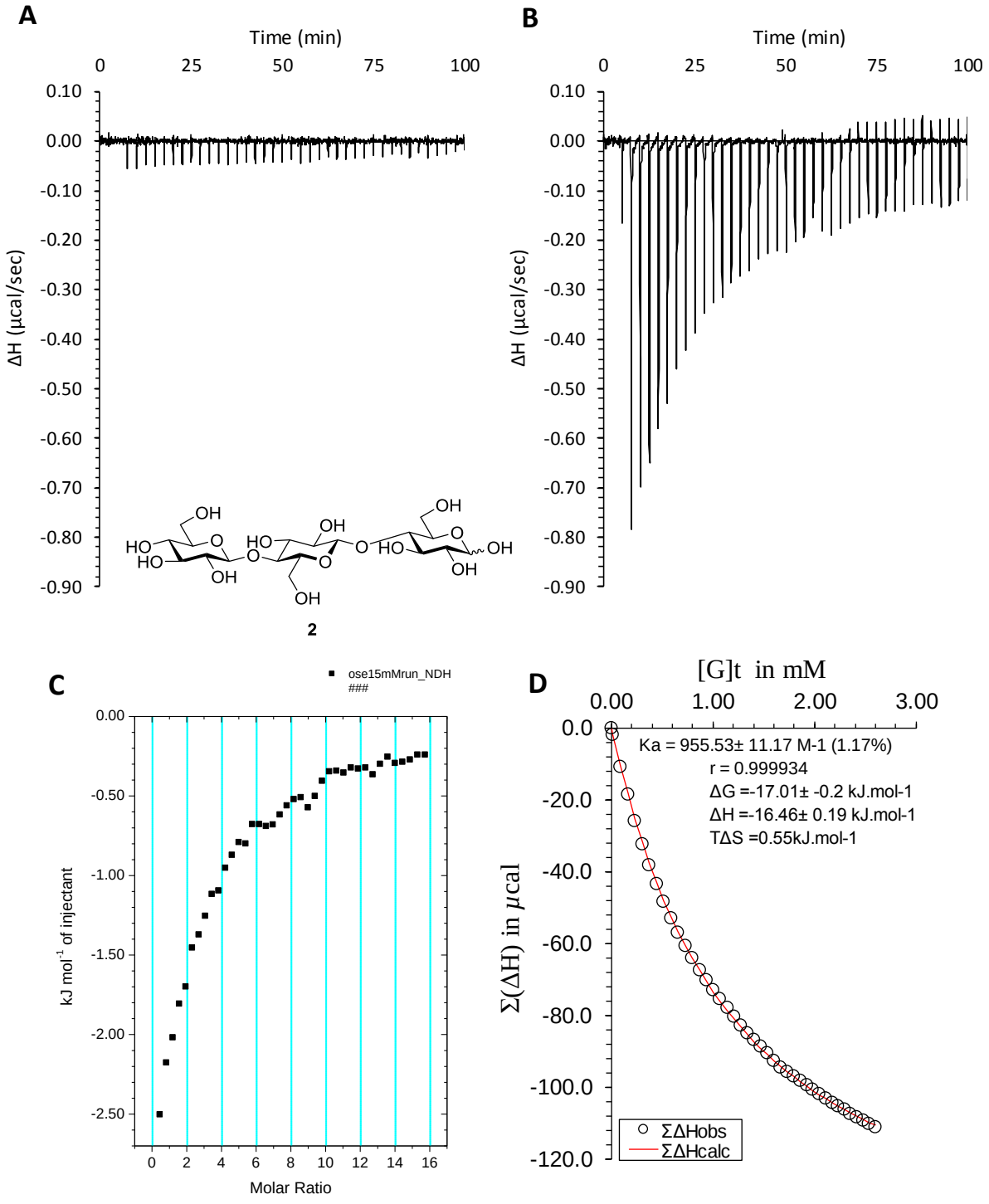


Figure 117 ITC binding results for **1** (0.2 mM) titrated with D-cellotriose **2** (15 mM) in water at 298 K, in which: A) shows the blank run (addition of substrate into water); B) shows the titration (substrate into receptor); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 955 ± 11 M^{-1}).

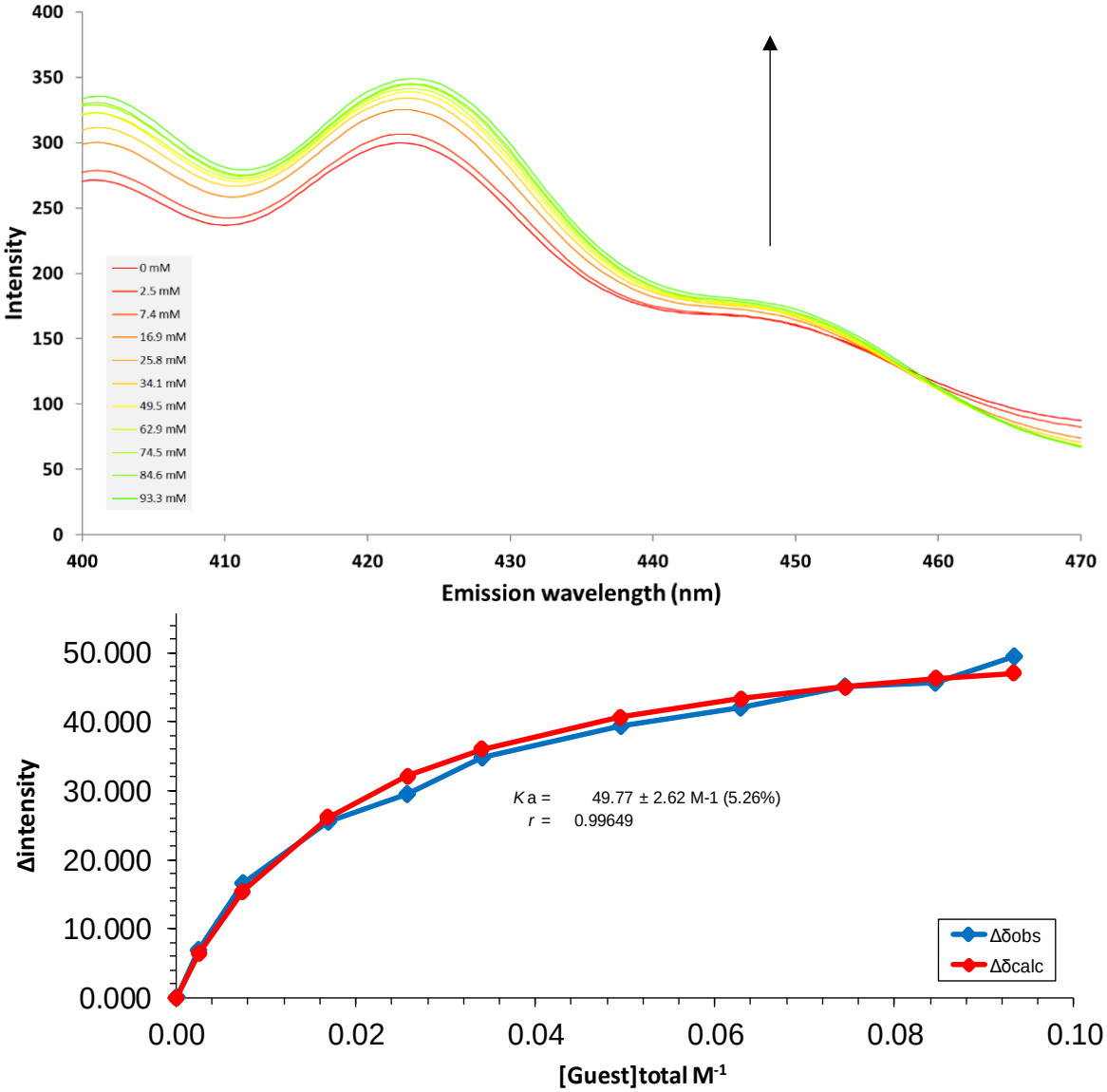


Figure 118 Fluorescence titration results for D-cellobiose **31** (150 mM) and receptor **1** (10 µM) in water (pH 7.4) at 298 K. Excitation wavelength of 395 nm used, while monitoring emission intensity at 423 nm. Changes in emission intensity (Δintensity) at 423 nm were plotted against concentration of D-cellobiose **31**. The calculated values for Δintensity are overlaid with the observed values, giving a K_a = 49.8 ± 2.6 M^{-1} (5.3%).

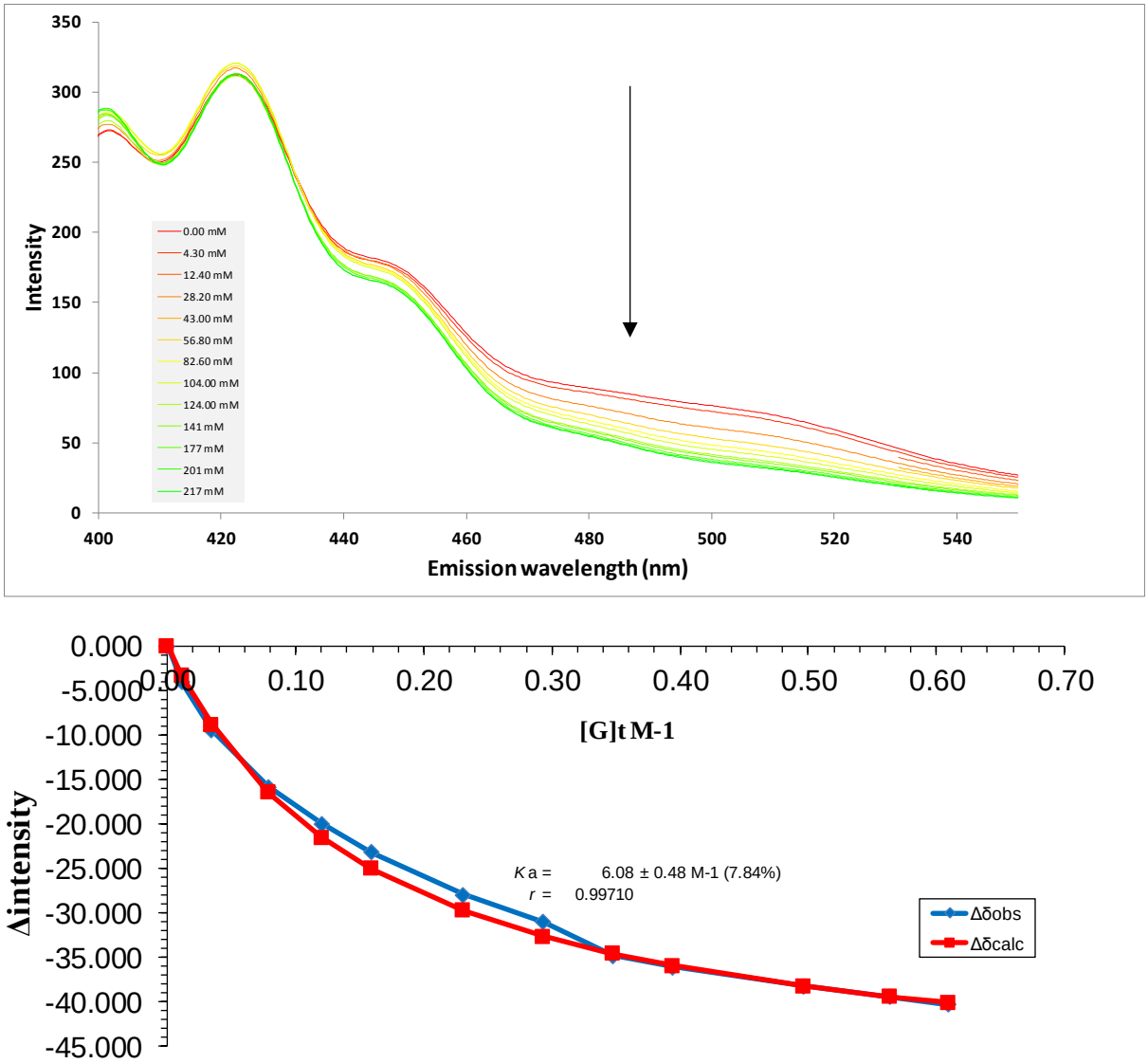


Figure 119 Fluorescence titration results for D-glucose **4** (700 mM) and receptor **1** (10 µM) in water (pH 7.4) at 298 K. Excitation wavelength of 395 nm used, while monitoring emission intensity at 480 nm. Changes in emission intensity (Δintensity) at 423 nm were plotted against concentration of D-glucose **4**. The calculated values for Δintensity are overlaid with the observed values, giving a K_a = 6.1 ± 0.5 M^{-1} (7.8%).

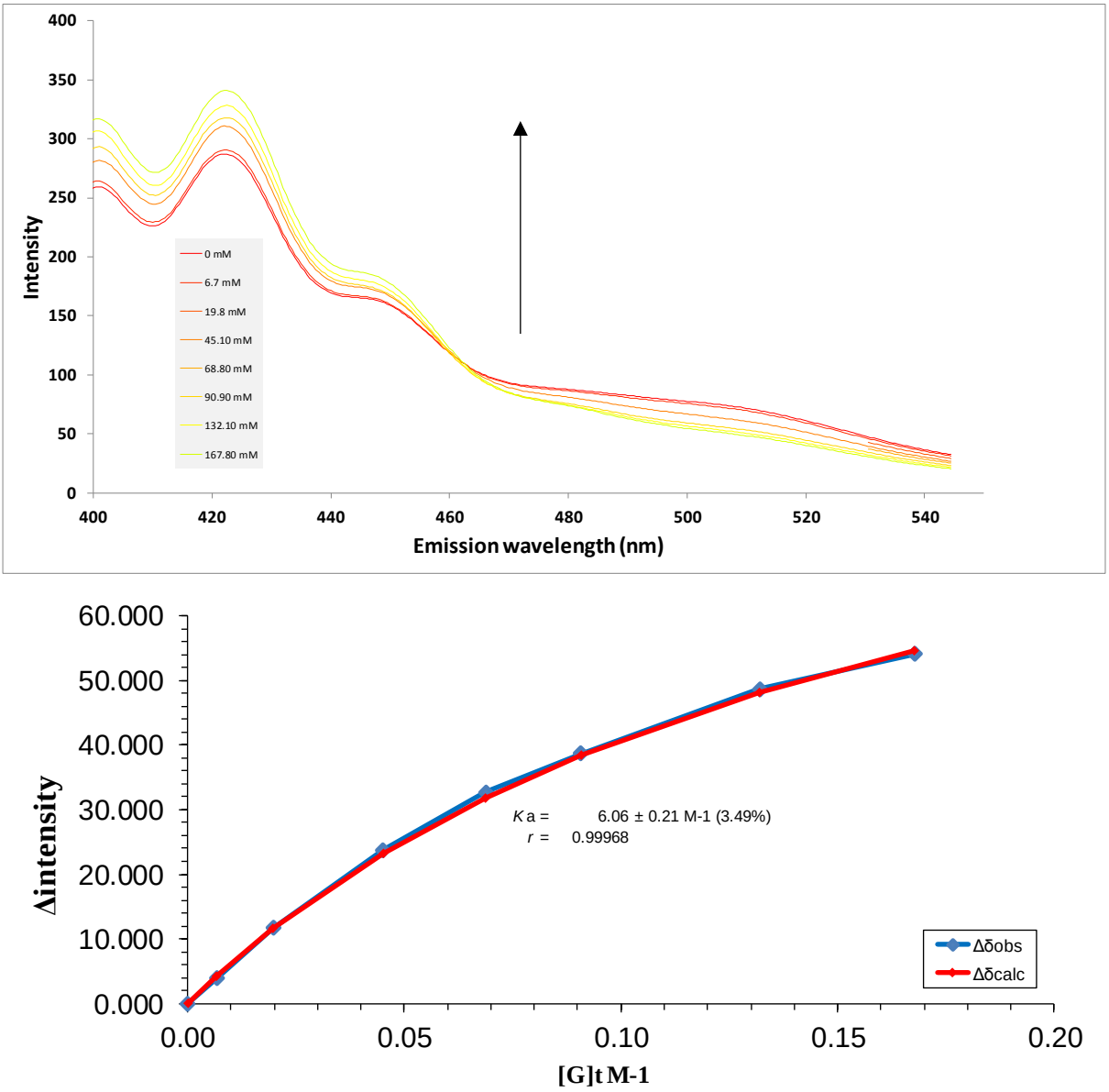


Figure 120 Fluorescence titration results for methyl β-D-glucoside **14** (400 mM) and receptor **1** (10 µM) in water (pH 7.4) at 298 K. Excitation wavelength of 395 nm used, while monitoring emission intensity at 423 nm. Changes in emission intensity (Δintensity) at 423 nm were plotted against concentration of D-glucose **4**. The calculated values for Δintensity are overlaid with the observed values, giving a K_a = 6.1 ± 0.2 M^{-1} (3.5%).

9.2 Triethylbenzene receptor (3)

9.2.1 Triethylbenzene receptor (3) studies

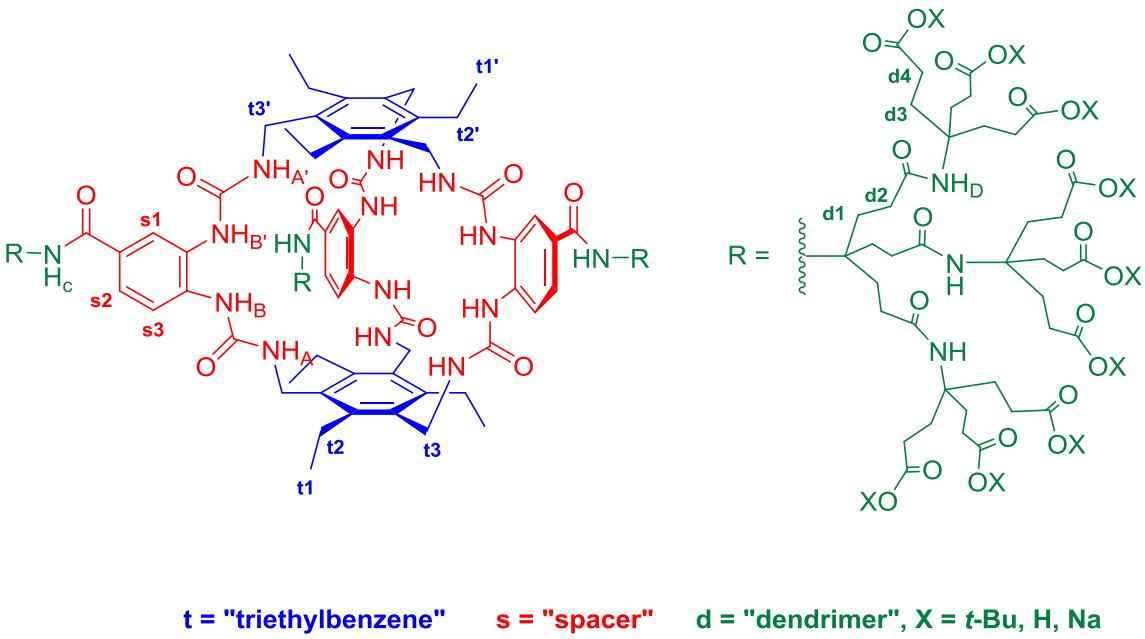


Figure 121 Numbering system used for the discussion of the conformational and binding properties of receptor **3**. Triethylbenzene proton designations (blue) are prefixed with "t" (i.e. t1, t1' etc.), spacer CH protons (red) are prefixed with "s" and protons from the dendritic solubilising groups are prefixed with "d". X can be H, Na, or $C(CH_3)_3$.

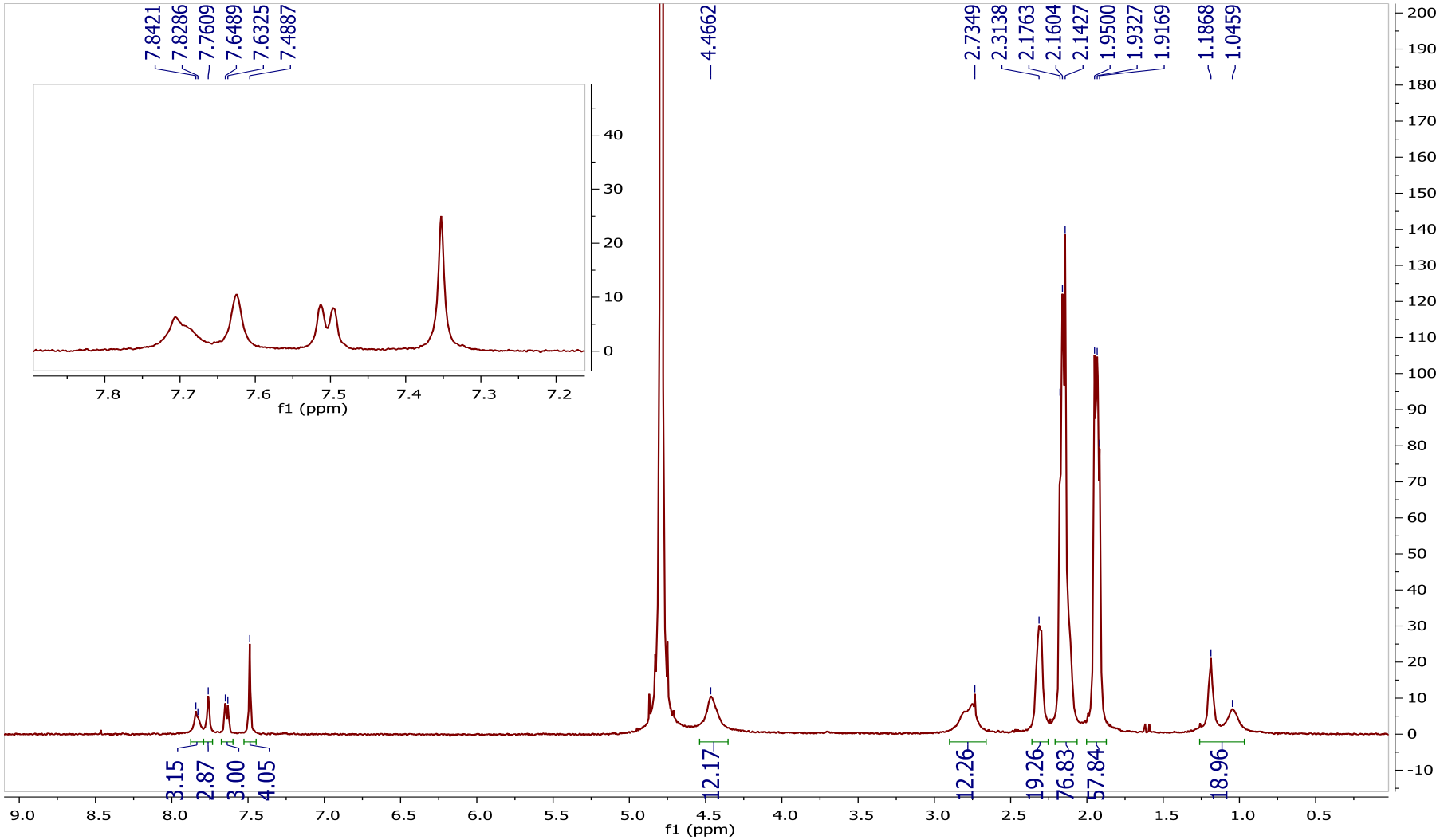


Figure 122 ^{1}H NMR spectrum of receptor **3** (0.5 mM) in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K.

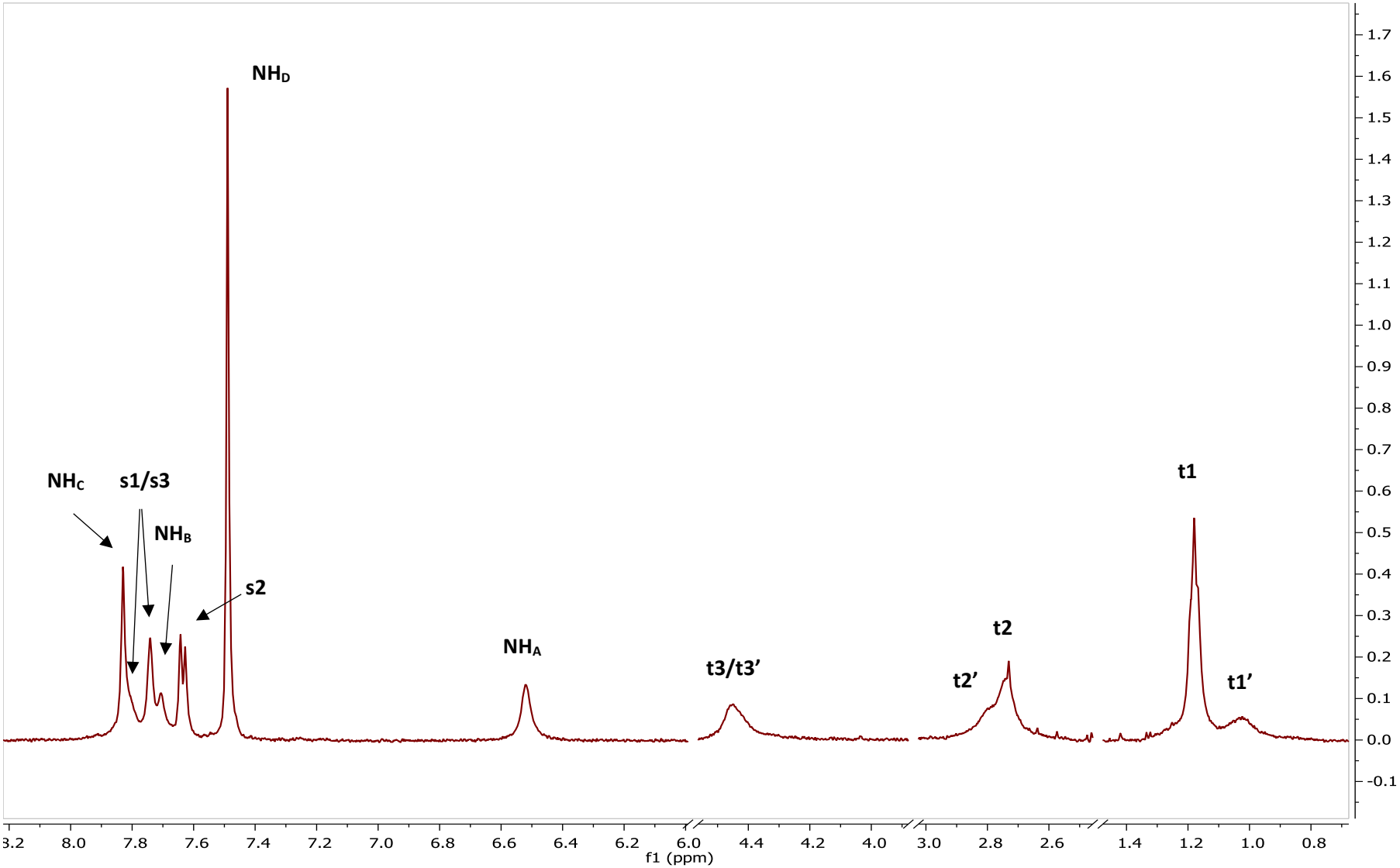


Figure 123 Partial ^{1}H NMR spectra, with assignments, of receptor **3** (0.5 mM) in D_2O/H_2O (1:9) with 10 mM phosphate buffer (pH 7.4) at 298 K.

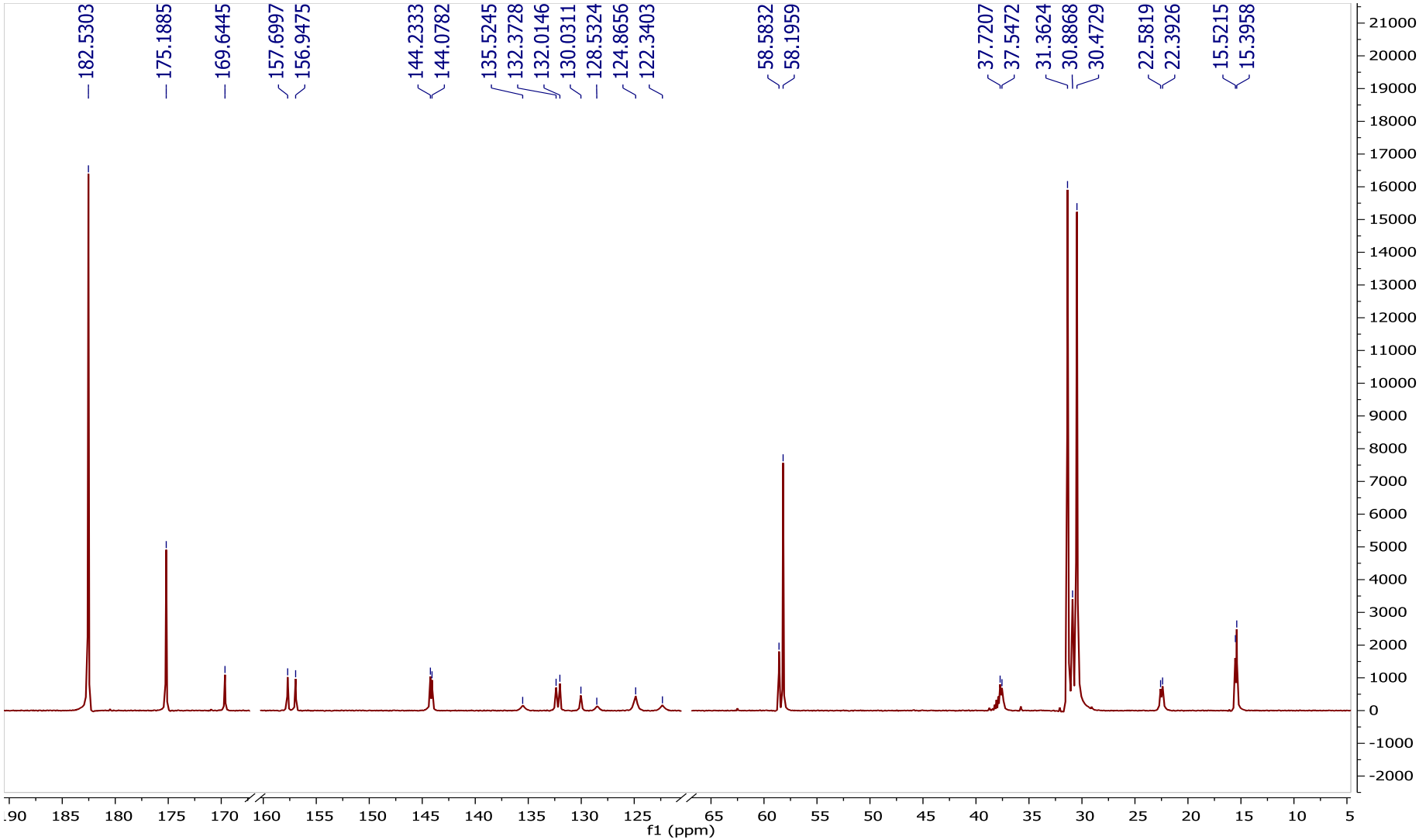


Figure 124 Partial ^{13}C NMR spectra of receptor **3** (10 mM) in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K.

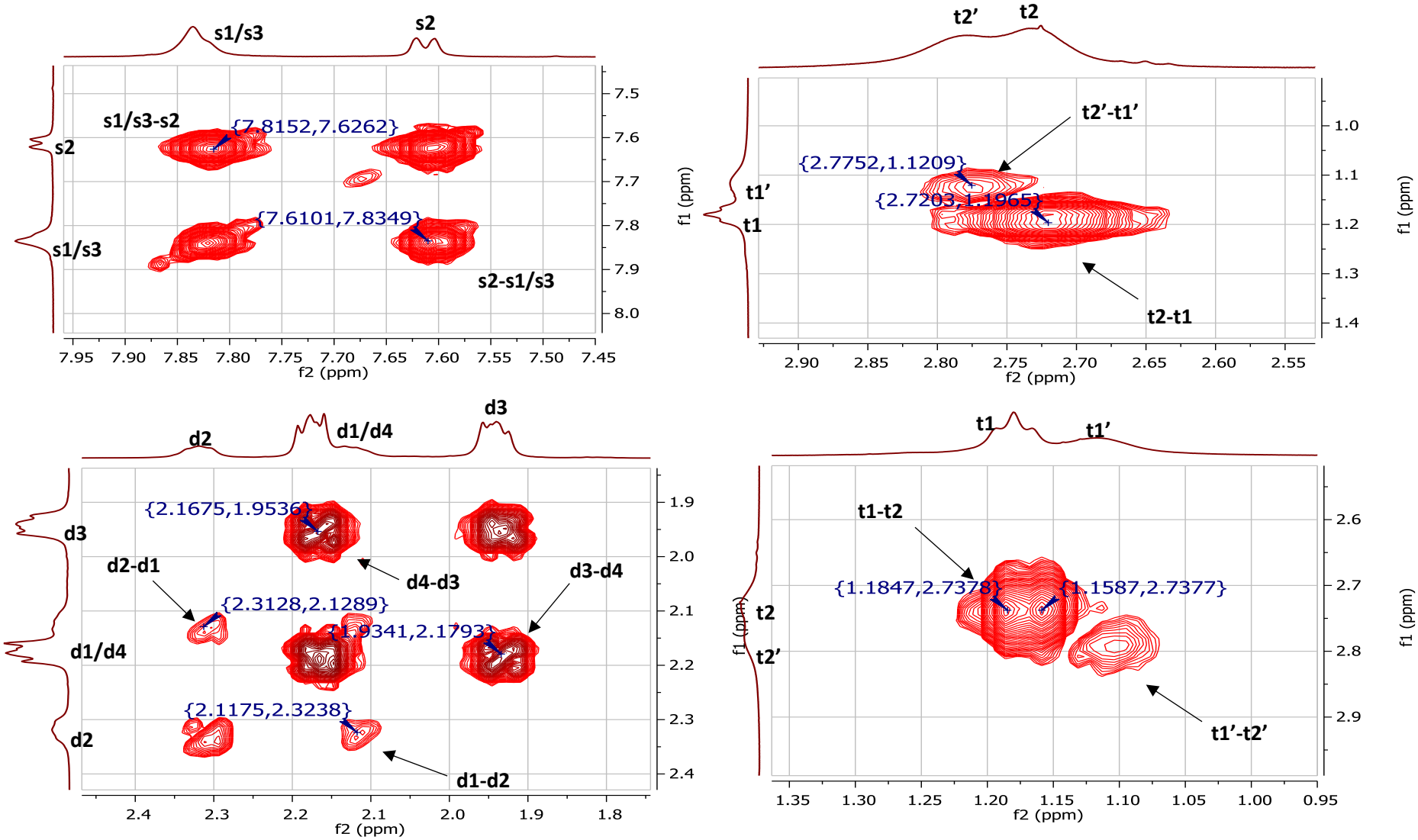


Figure 125 Partial (^{1}H-^{1}H correlation) COSY NMR spectra for receptor **3** (2mM) in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K.

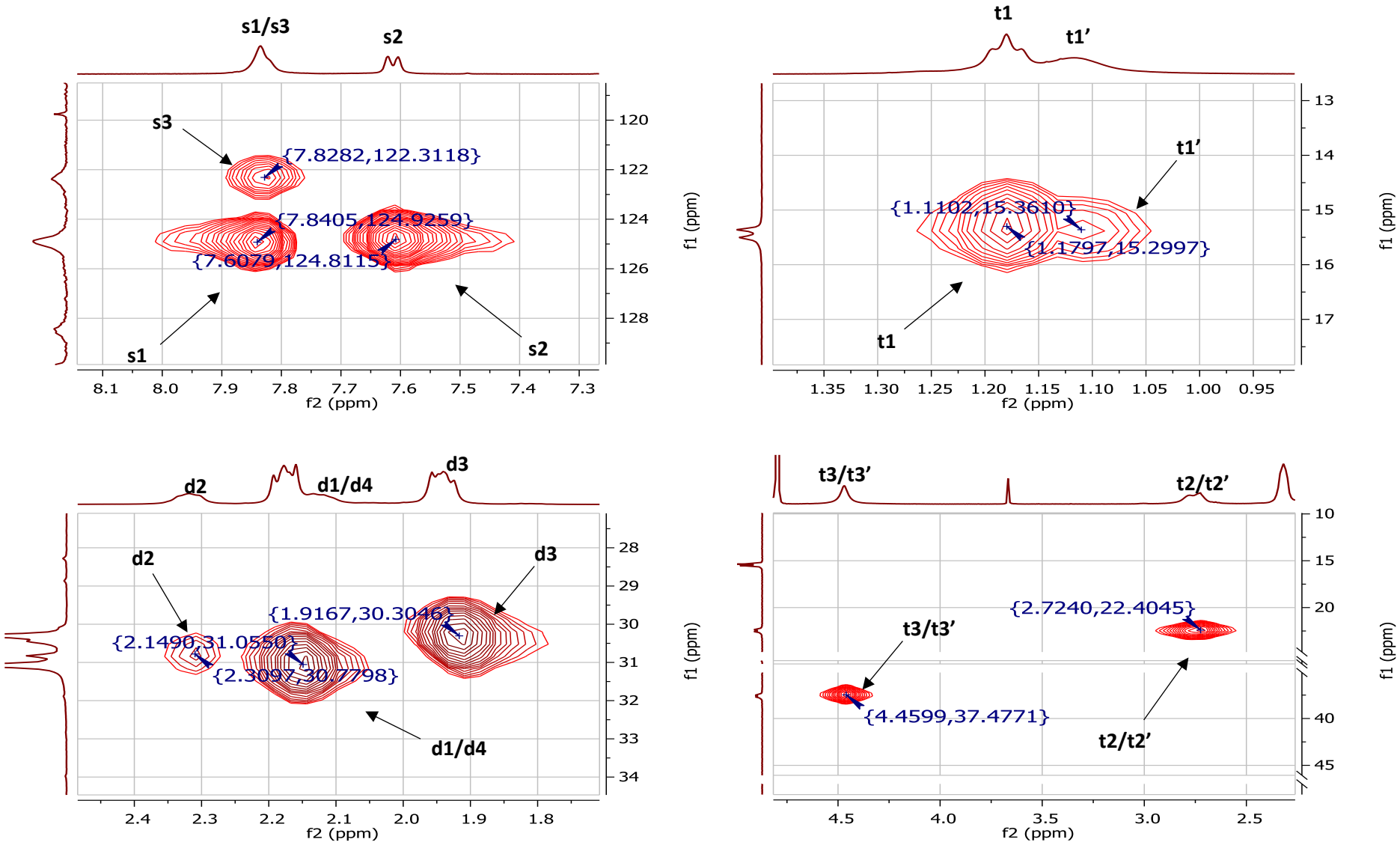


Figure 126 Partial (^{1}H-^{13}C correlation, one bond) HSQC NMR spectra for receptor **3** (10 mM) in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

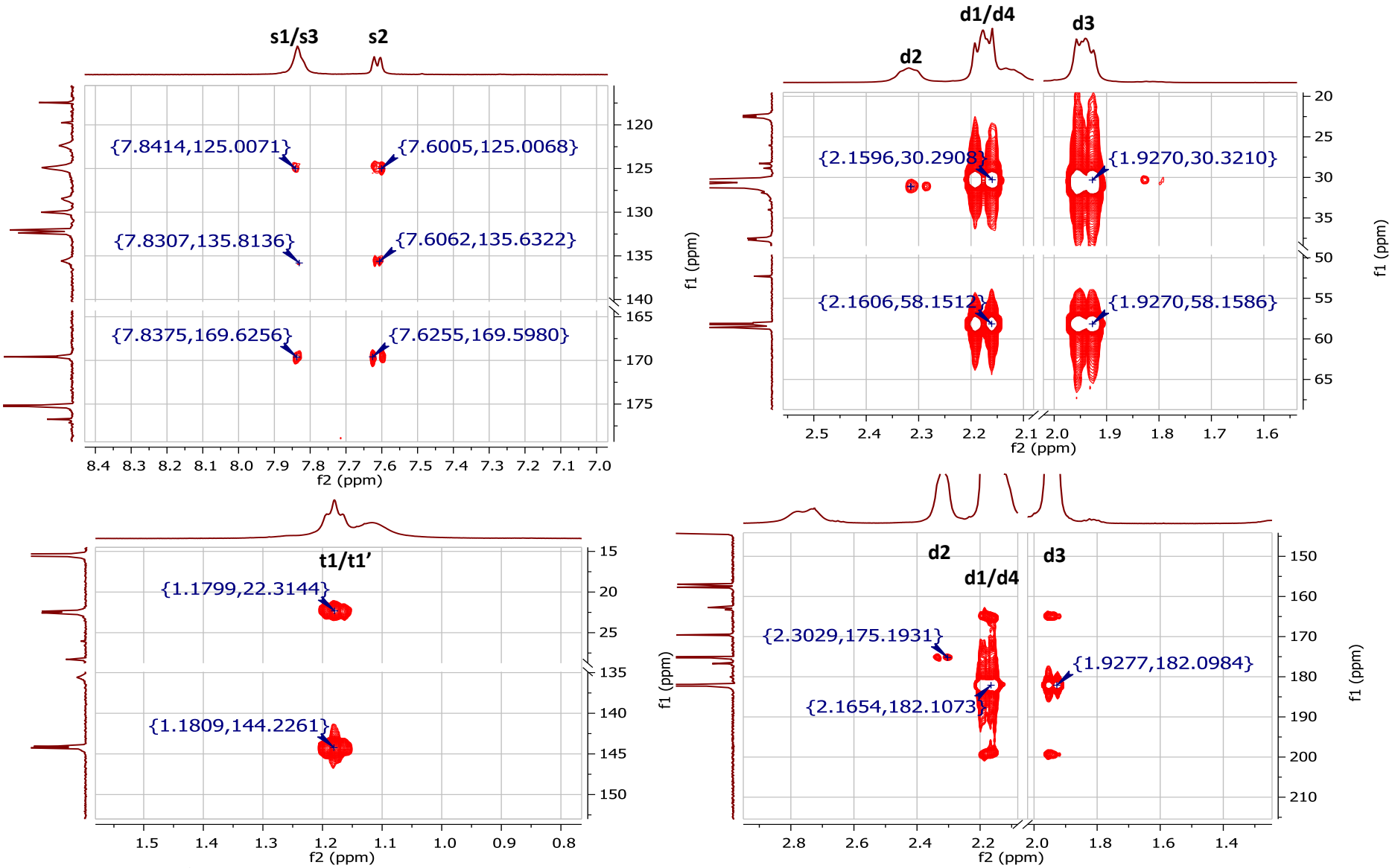


Figure 127 Partial (^{1}H-^{13}C correlation, multiple bond) HMBC NMR spectra for receptor **3** (10 mM) in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

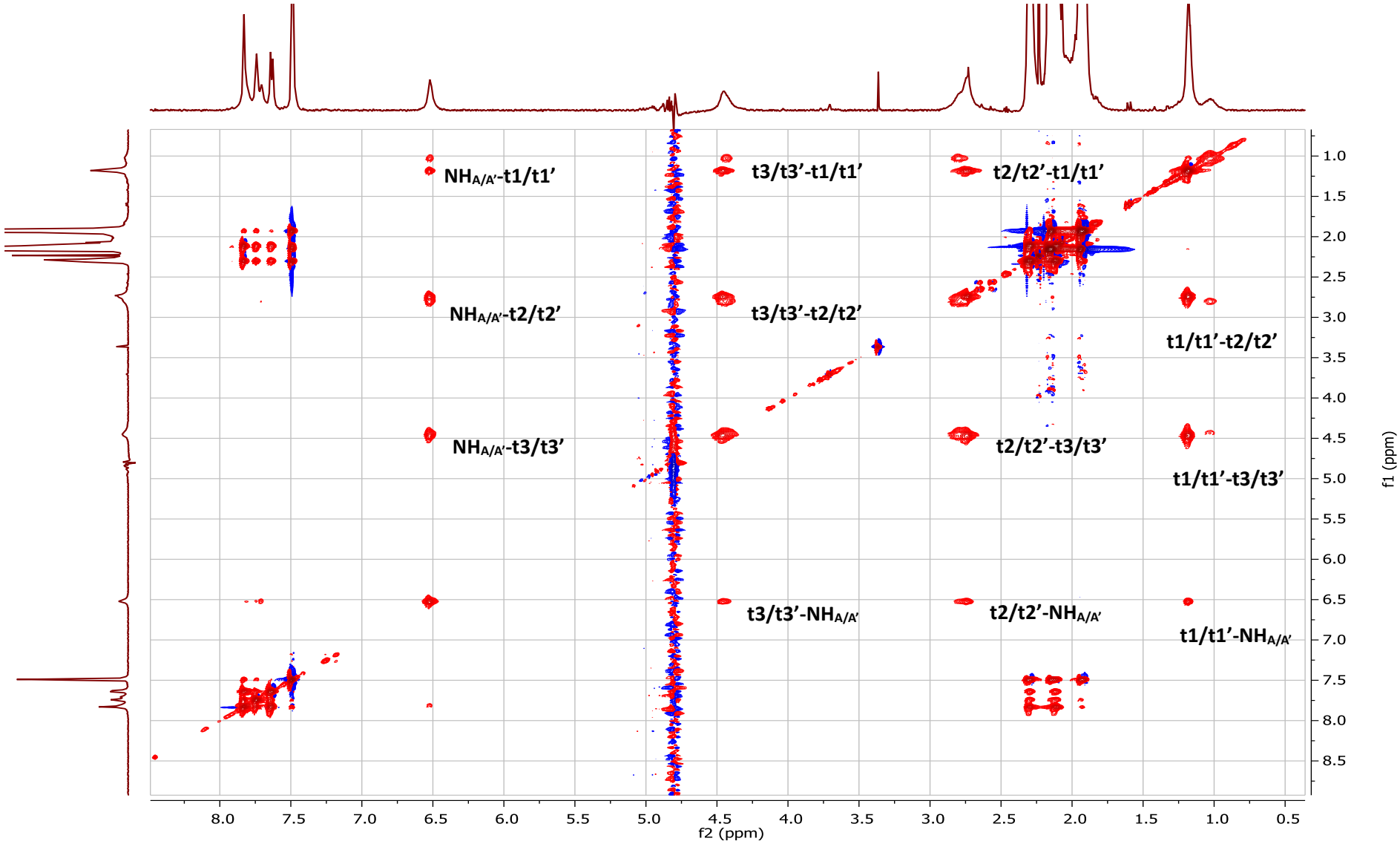


Figure 128 (^{1}H-^{1}H correlation, through space) NOESY NMR spectrum of receptor **3** (1 mM) in 9:1 H_2O/D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

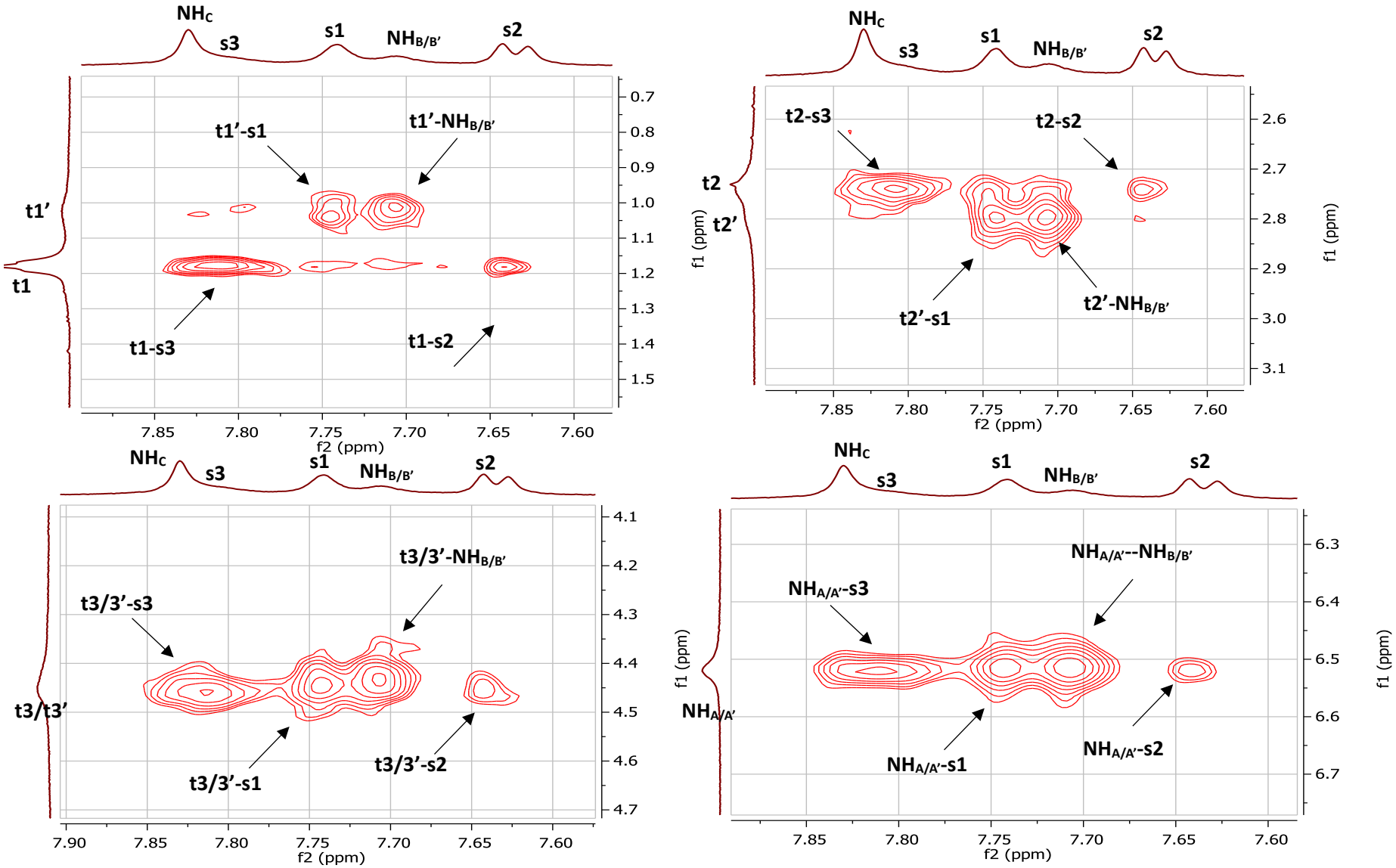


Figure 129 Partial (^{1}H-^{1}H correlation, through space) NOESY NMR spectra for receptor **3** (1 mM) in 9:1 H_2O/D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

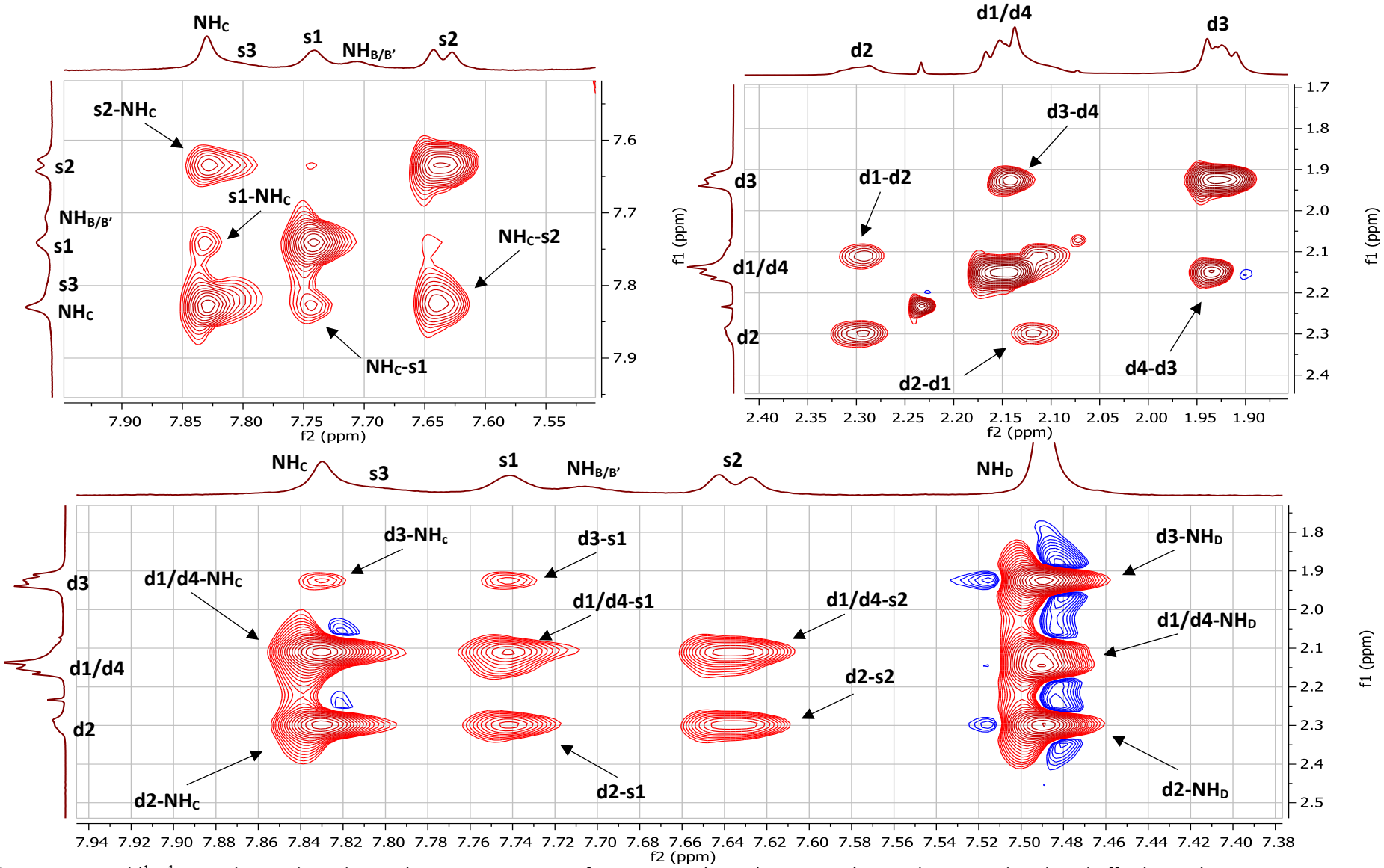


Figure 130 Partial (^{1}H-^{1}H correlation, through space) NOESY NMR spectra for receptor **3** (1 mM) in 9:1 H_2O/D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K.

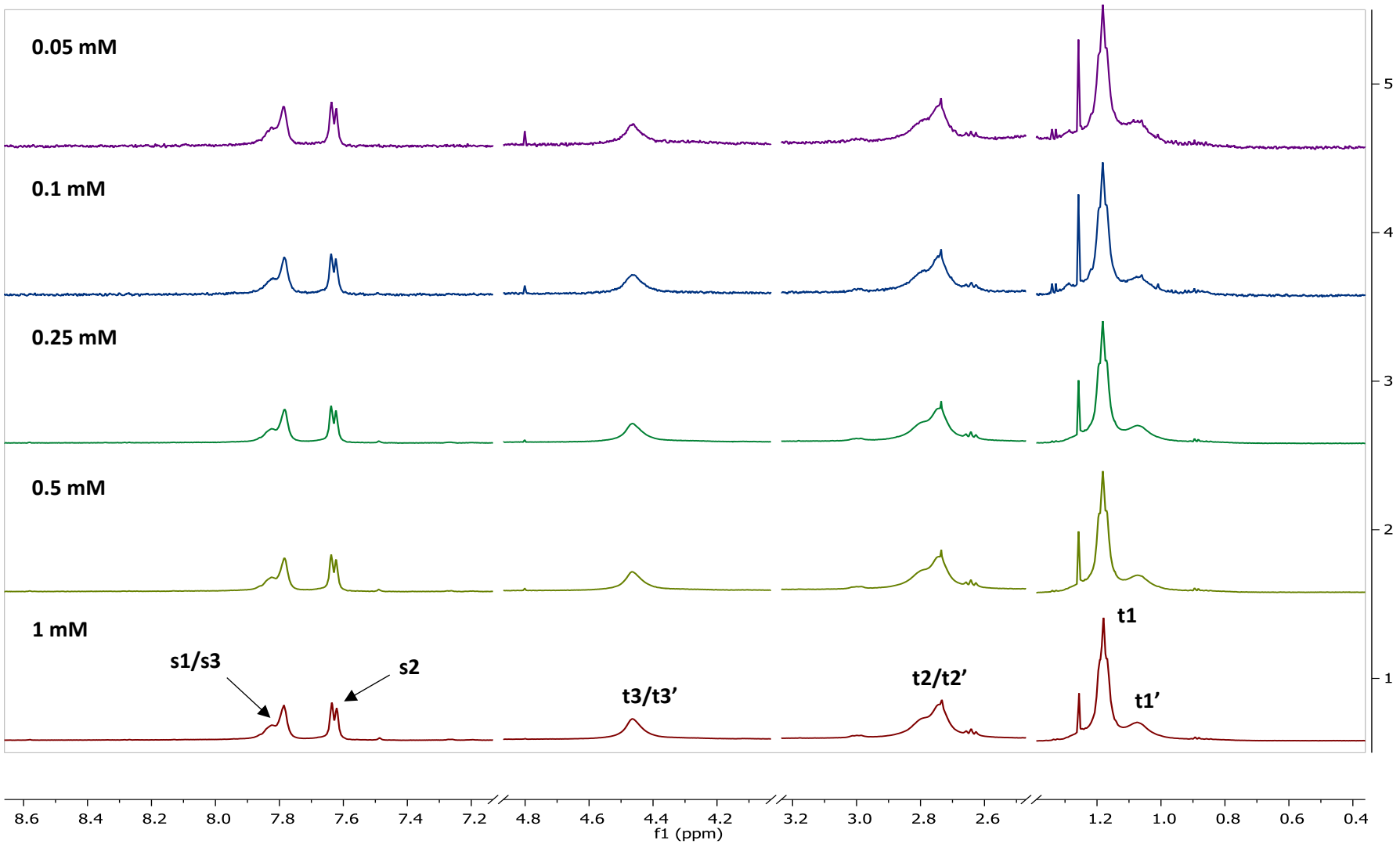


Figure 131 Partial ^{1}H NMR spectra showing receptor **3** at various concentrations in D_2O with 10 mM phosphate buffer (pH 7.4). No change in spectra across concentration range (1 mM – 0.05 mM) suggests receptor **3** is monomeric at these concentrations.

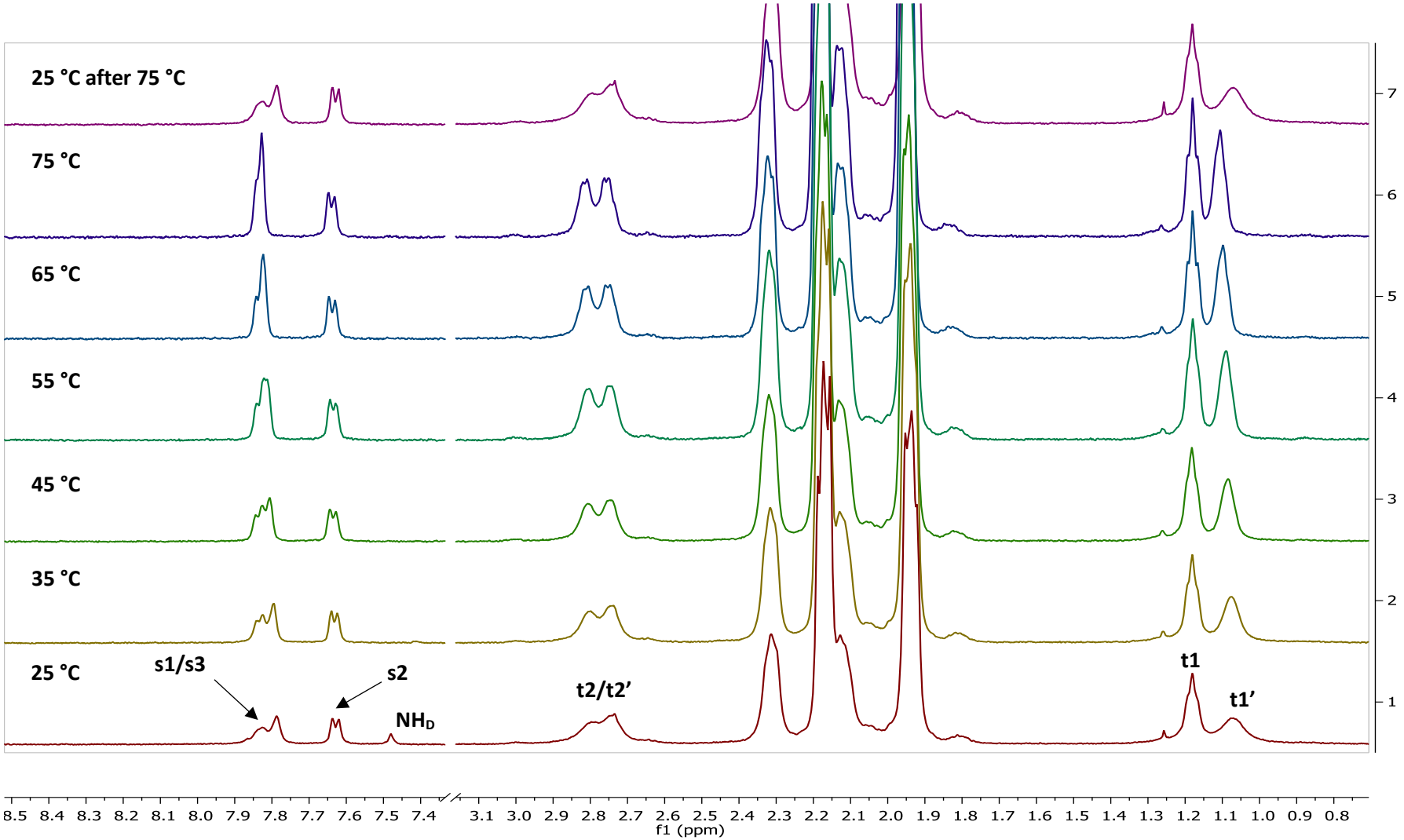


Figure 132 Partial ^{1}H NMR spectra showing receptor **3** at various temperatures in D_2O with 10 mM phosphate buffer (pH 7.4). Chemical shifts (δ, ppm) relative to proton t1 (1.18 ppm). Receptor signals sharpen upon heating, most notably t1/t1' and t2/t2' and s3. Signal for proton t3/t3' is omitted for clarity due to overlap with the residual solvent peak that varied in chemical shift upon heating. NH_D disappears over time due to deuterium exchange

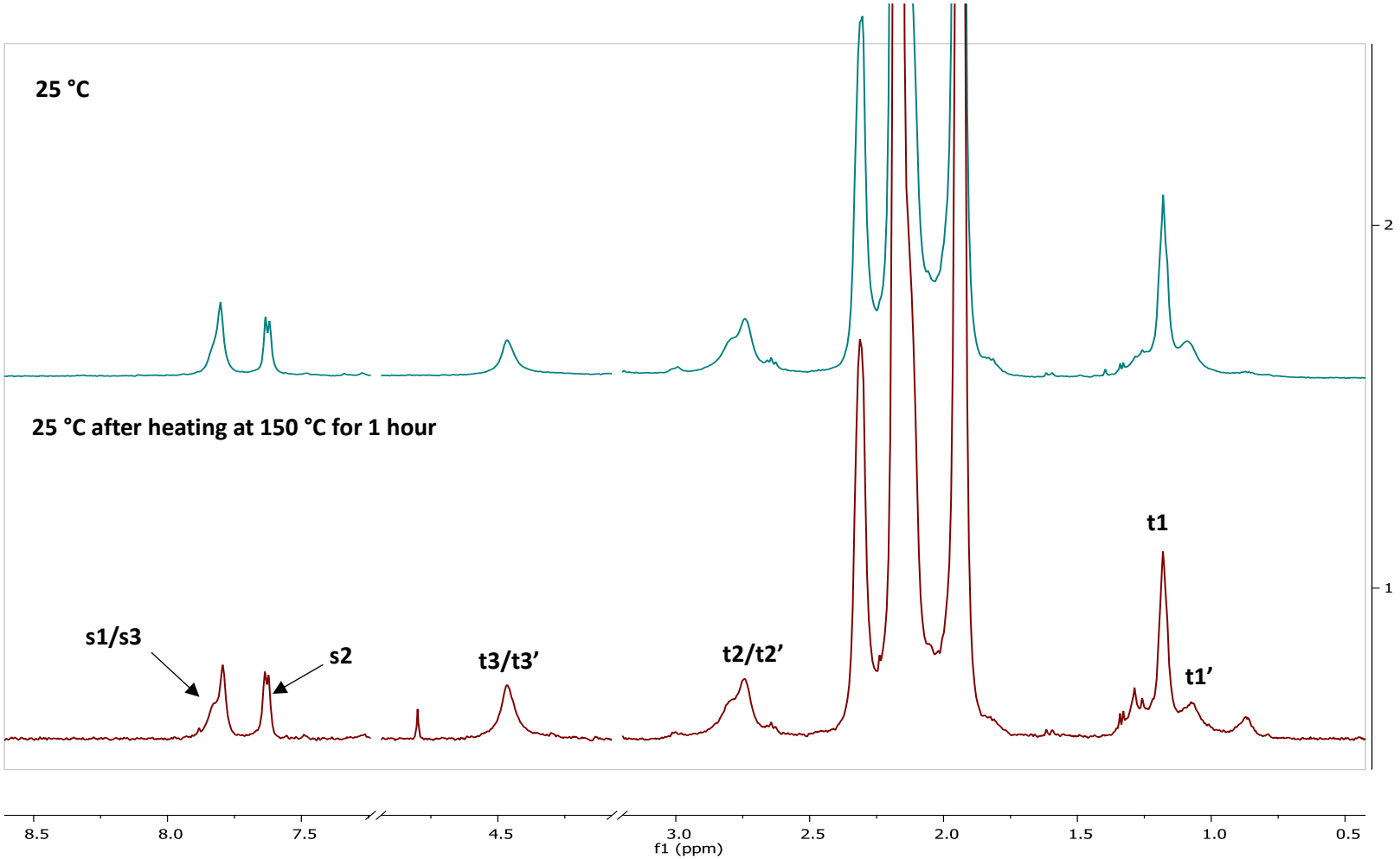


Figure 133 Partial ^{1}H NMR spectra showing receptor **3** (0.1 mM) at 25 °C (top) and receptor **3** (0.1 mM) 25 °C after heating at 150 °C for 1 hour in the solid state (bottom), showing no signs of decomposition. Both spectra recorded in D_2O with 10 mM phosphate buffer (pH 7.4).

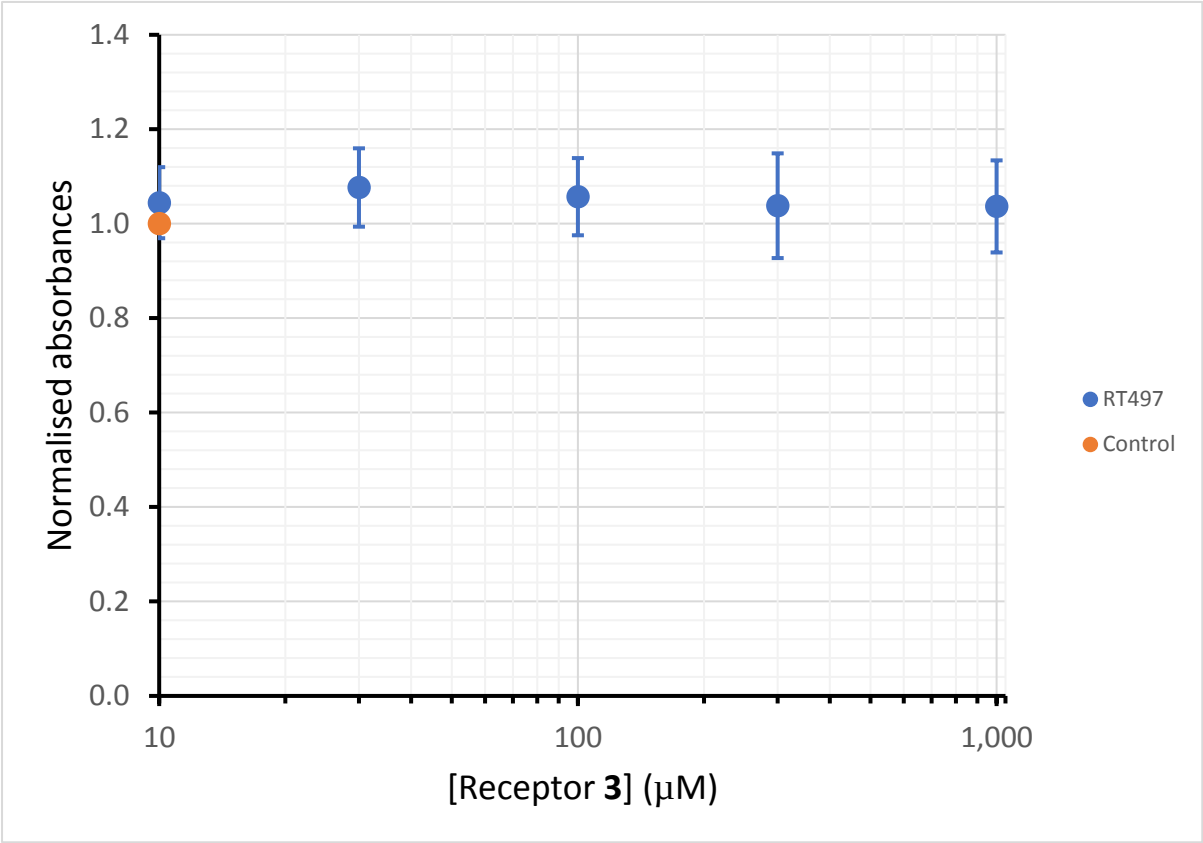
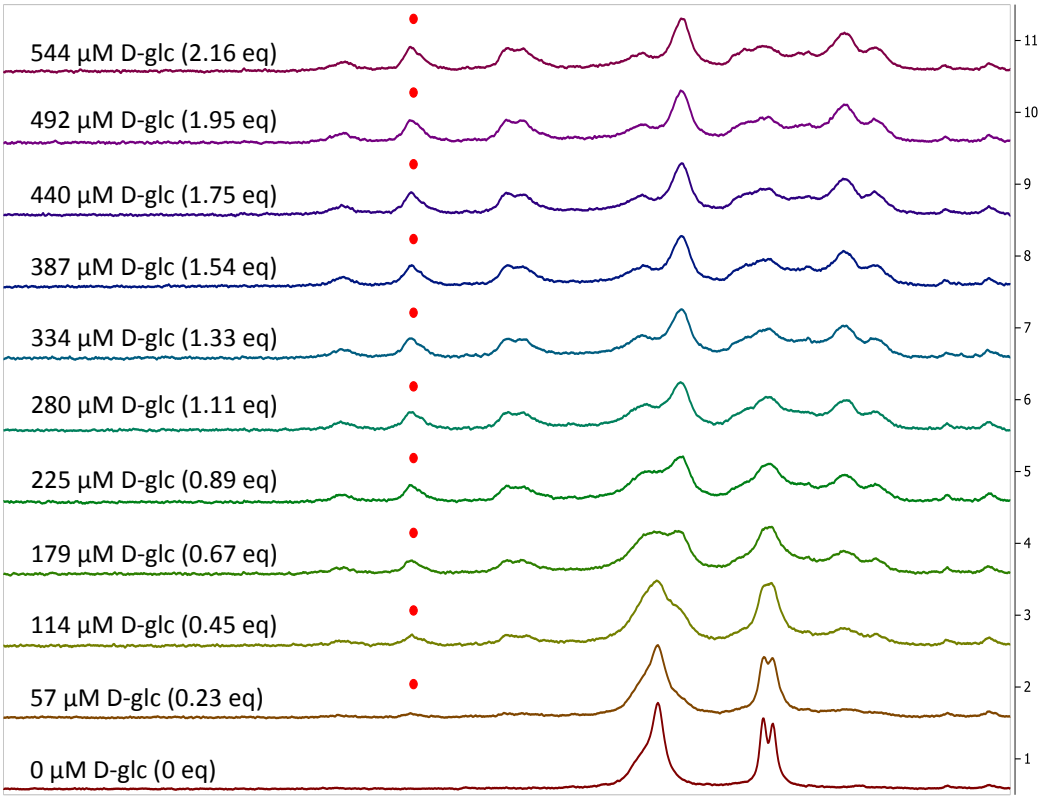


Figure 134 HeLa cells were incubated with normal cultural medium (control) or medium containing **3** at indicated concentrations (10 -1000 µM, dissolved in cultural medium) for 18 hours. Cell viability was assessed using sodium 2,3,-bis(2-methoxy-4-nitro-5-sulfophenyl)-5-[(phenylamino)-carbonyl]-2H-tetrazolium) (XTT). Cell viability is presented as specific absorbance values normalized to that of control (means ± SEM; n = 20-28 with 4 replicates per treatment). There is no significant difference between the cells treated with normal cultural medium (orange) and those treated with medium with **3** incorporated at all tested concentrations (blue). These data suggest that receptor **3** at concentrations as high as 1000 µM is without toxic effect. Study performed by Dr Hongyu Li.

9.2.2 Binding studies for triethylbenzene receptor (3) and D-glucose (4)

.



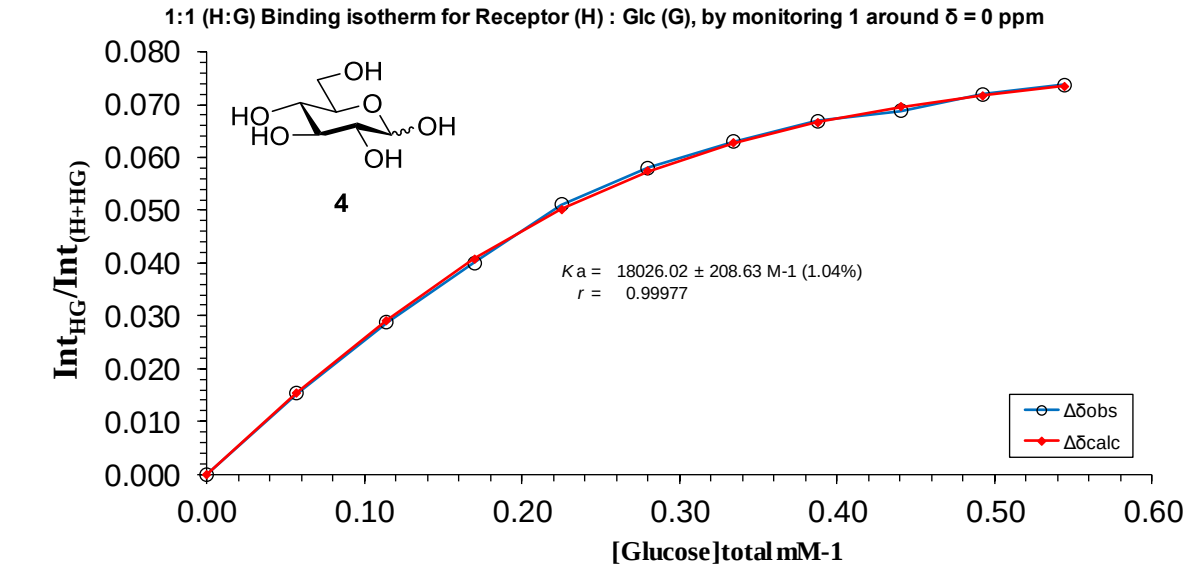


Figure 135 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.25 mM) titrated with a combined solution of D-glucose **4** (9.6 mM) and receptor **3** (0.25 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with slow exchange on NMR timescale. Integrations of the peak at 8.19 ppm (denoted with •) versus the region 8.35-7.39 ppm were plotted against D-glucose **4** concentration (mM). The calculated values for the integrals are overlaid with the observed values, giving K_a = 18,026 ± 208 M^{-1} (1.04%).

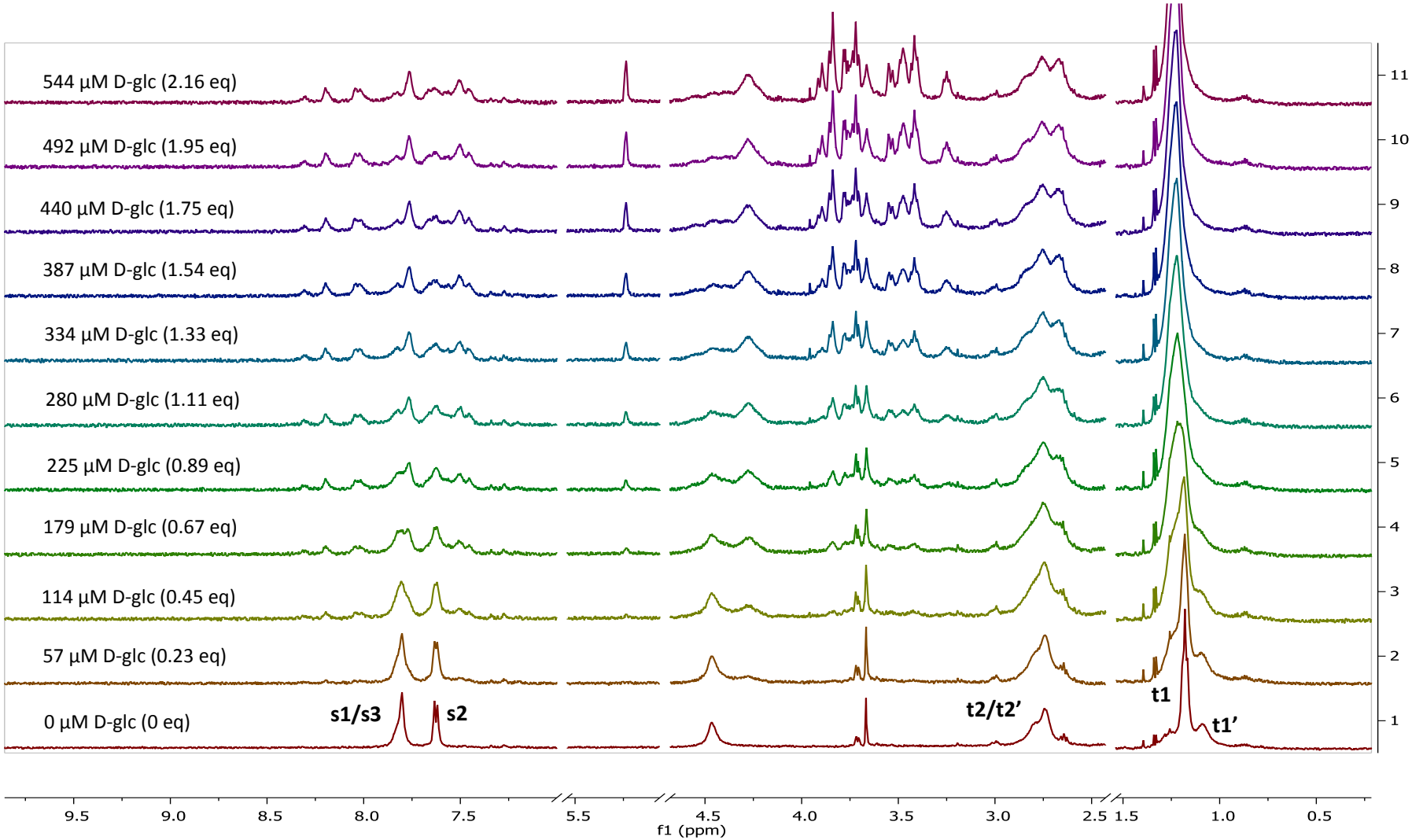


Figure 136 ^{1}H NMR spectra for receptor **3** (0.25 mM) titrated with a combined solution of D-glucose **4** (9.6 mM) and receptor **3** (0.25 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K.

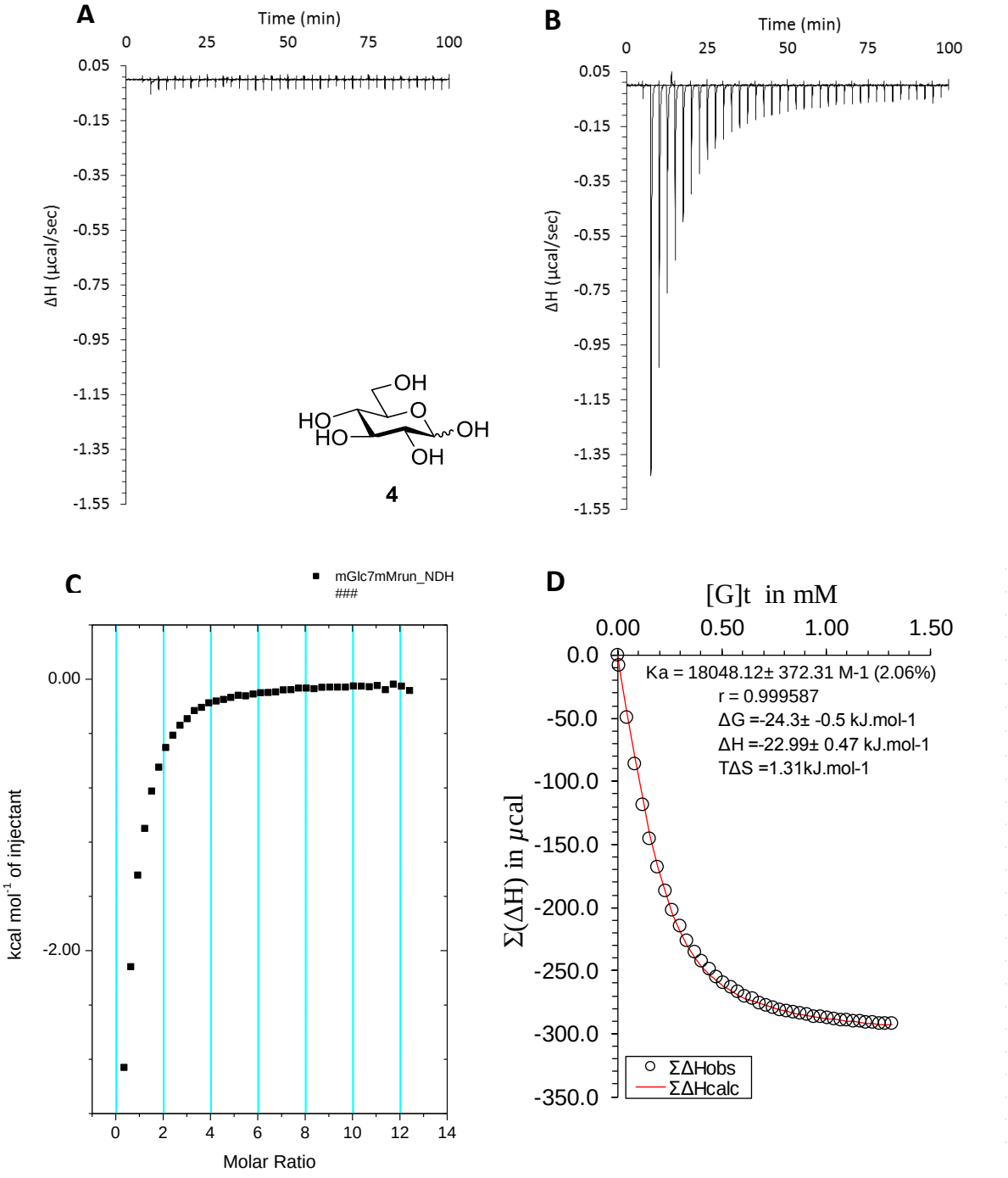


Figure 137 ITC binding results for receptor **3** (0.2 mM) titrated with D-glucose **4** (7.5 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio and D) shows the fit calculated using a 1:1 binding model (K_a = 18,000 ± 372 M^{-1}).

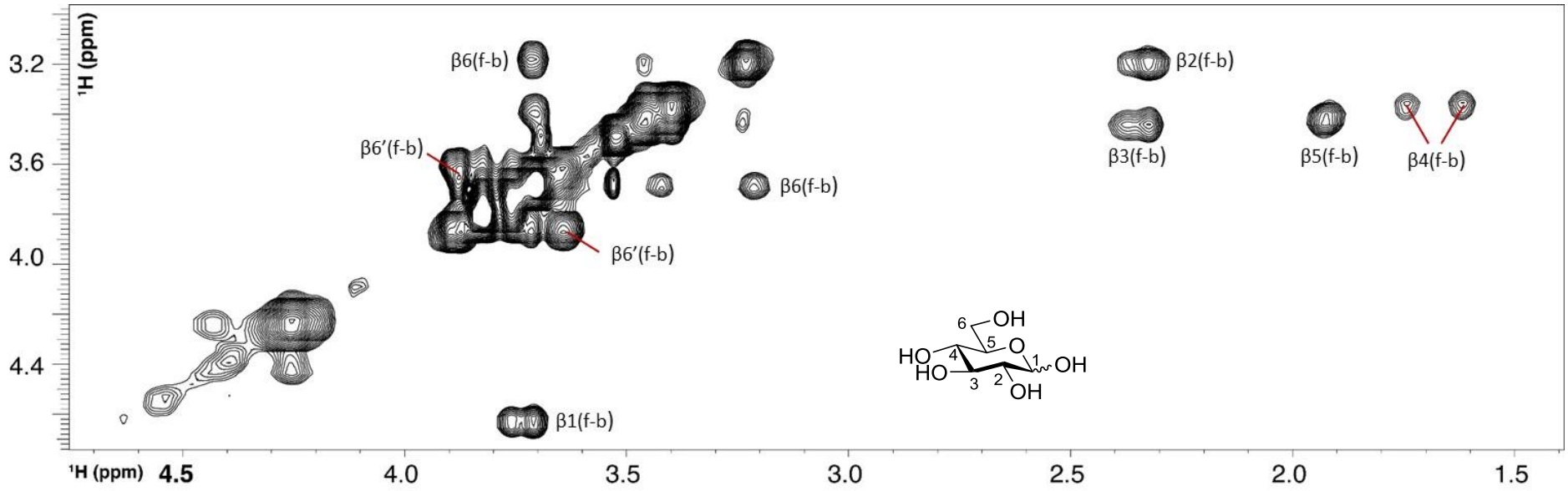


Figure 138 Partial ^{1}H-^{1}H NMR ROESY spectrum of receptor **3** (2 mM) with D-glucose **4** (5 mM, 2.5 equivalents) in D_2O at 298 K. Chemical exchange peaks are observed between the D-glucose bound to the receptor (whose signals appear more upfield) and free D-glucose. These exchange signals are labelled accordingly (i.e. β2(f-b) indicates chemical exchange between free ("f") and bound ("b") β-D-glucose for the proton at position 2). Exchange peaks are observed for only the β anomer under these conditions (mixing time of 100 ms). Exchange peaks for the α anomer were observed at much longer mixing times (>500 ms). Quantification of the selectivity of **3** for the α and β anomers could not be achieved using this method, however determination of selectivity was achieved using ^{1}H-^{13}C HSQC (see Figure 139**Error! Reference source not found.** below). Comparison of the chemical shifts between the free protons for D-glucose allowed assignment of the ^{1}H NMR chemical shifts for the bound forms of β-D-glucose, which are listed in Table 10 below.

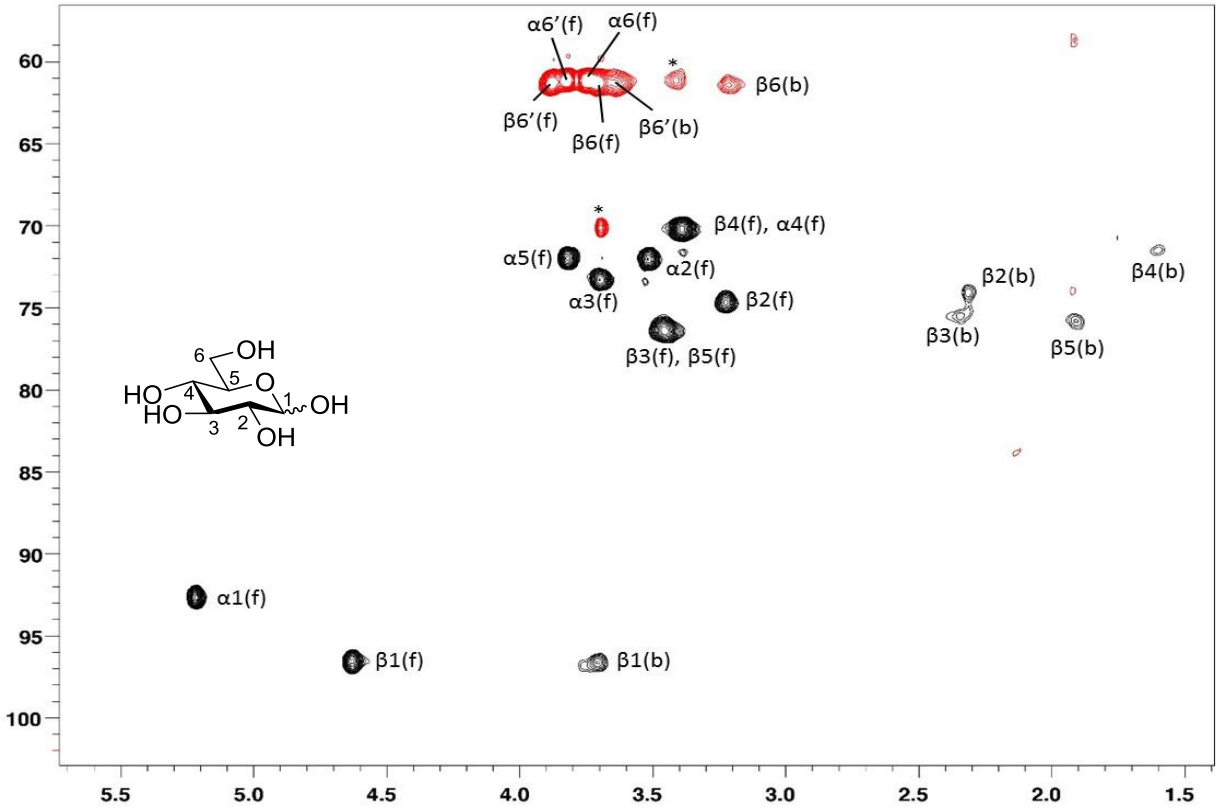
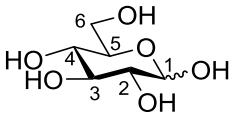


Figure 139 Partial (^{1}H-^{13}C correlation, one bond) HSQC NMR spectrum of receptor **3** (2 mM) with D-glucose **4** (5 mM, 2.5 equivalents) in D_2O at 298 K. NMR signals for D-glucose bound to the receptor (whose signals appear more upfield in the ^{1}H (f1) dimension, horizontal axis) and free glucose in solution. These signals for "free" and "bound" states are shown in black for CH protons and red for CH_2 protons, and are labelled accordingly (i.e. β2(f) indicates free β-D-glucose and β2(b) indicates the bound form for the proton at position 2). NMR signals corresponding to bound D-glucose are observed for the β anomer only, indicating an expected preference for binding the β anomer. The ratio of intensity of signals for bound β-D-glucose and baseline noise is measured as 6:1 (signal:noise). Assuming the bound signals for the α anomer are at most the same intensity as the baseline noise, then it can be assumed the selectivity for β and α-D-glucose is also at least 6:1 (β:α) but could potentially be higher.

Table 10 Chemical shifts (δ in ppm) of β-D-glucose **4** unbound and bound to receptor **3**. Values obtained in D_2O/H_2O (1:9) at 298 K. The values for difference in chemical shift (Δδ) upon complexation of glucose with receptor 2 are also given, indicating how shielded the CH protons of the carbon skeleton become when bound to the receptor.

β-D-glucose chemical shifts (δ, ppm)

H	Unbound	Bound to **2**	Δδ on binding to **2**
1	4.63	3.72	-0.91
2	3.23	2.34	-0.89
3	3.47	2.37	-1.10
4	3.39	1.63, 1.74	-1.76, -1.65
5	3.45	1.94	-1.51
6	3.71	3.23	-0.48
6’	3.88	3.64	-0.24



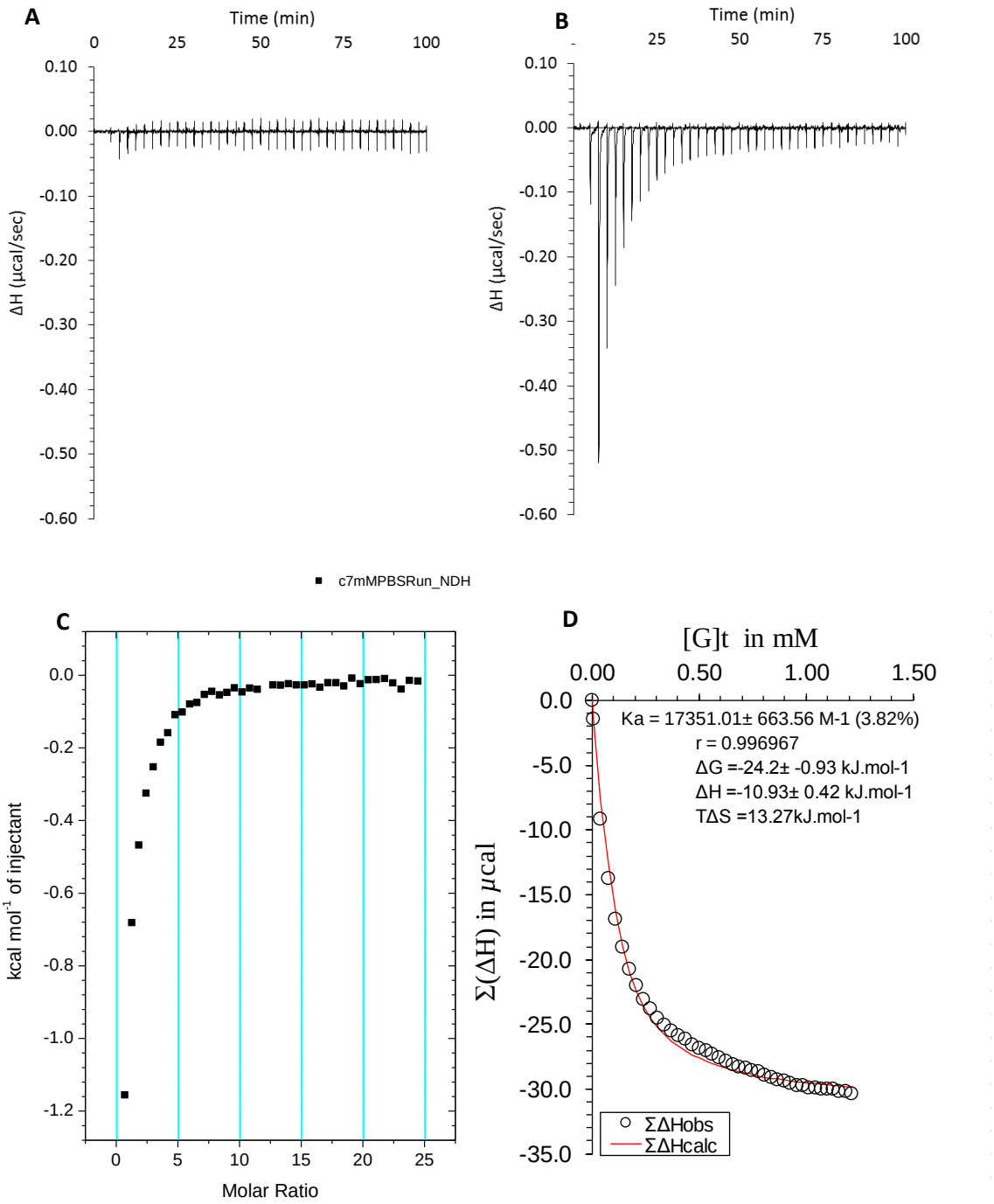


Figure 140 ITC binding results for receptor **3** (0.06 mM) titrated with D-glucose **4** (7 mM) in 10 mM phosphate buffered saline solution (pH 7), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 17,351 ± 663 M^{-1}).

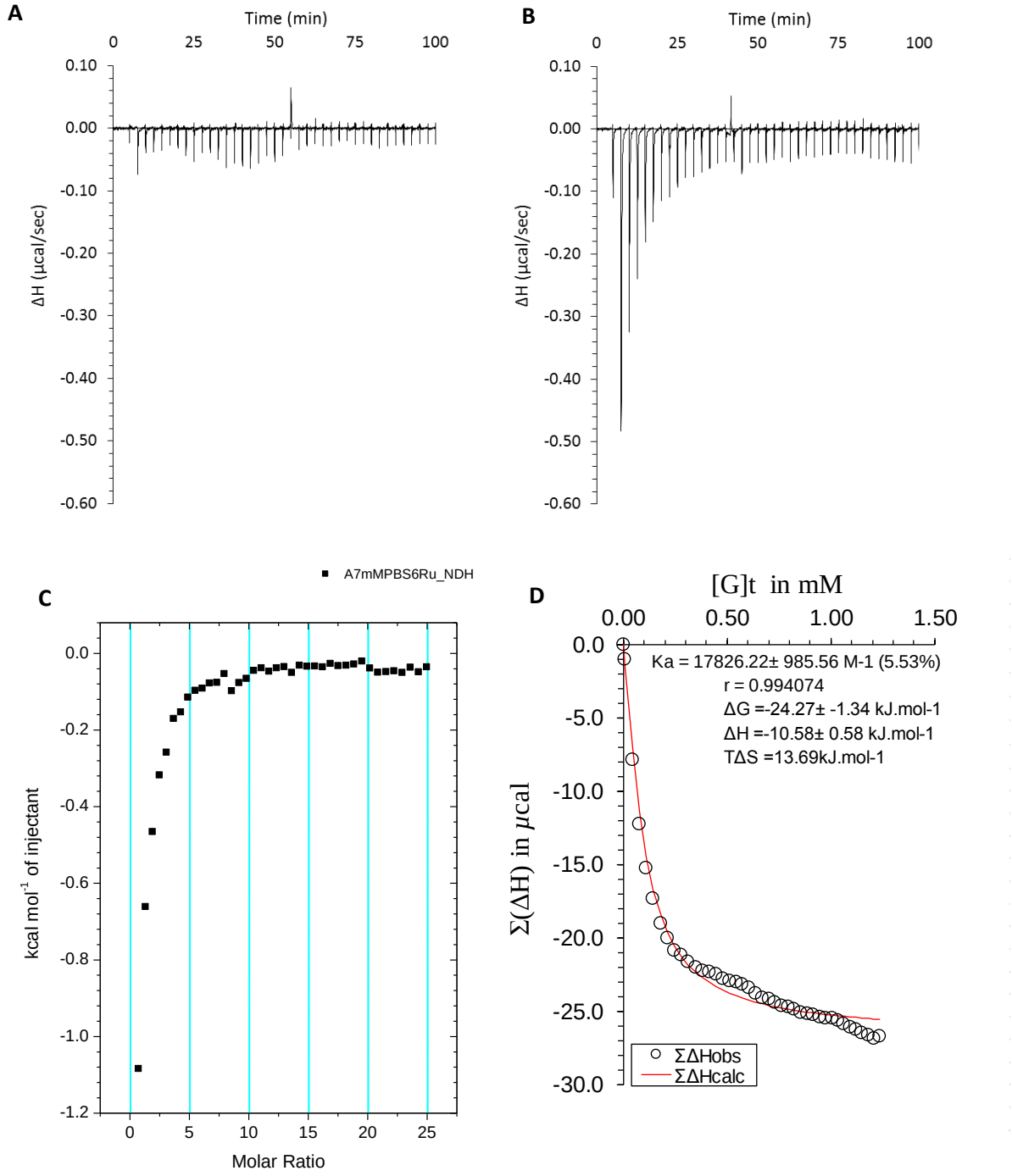


Figure 141 ITC binding results for receptor **3** (0.06 mM) titrated with D-glucose **4** (7 mM) in 10 mM phosphate buffered saline solution (PBS, pH 6), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 17,826 ± 985 M^{-1}).

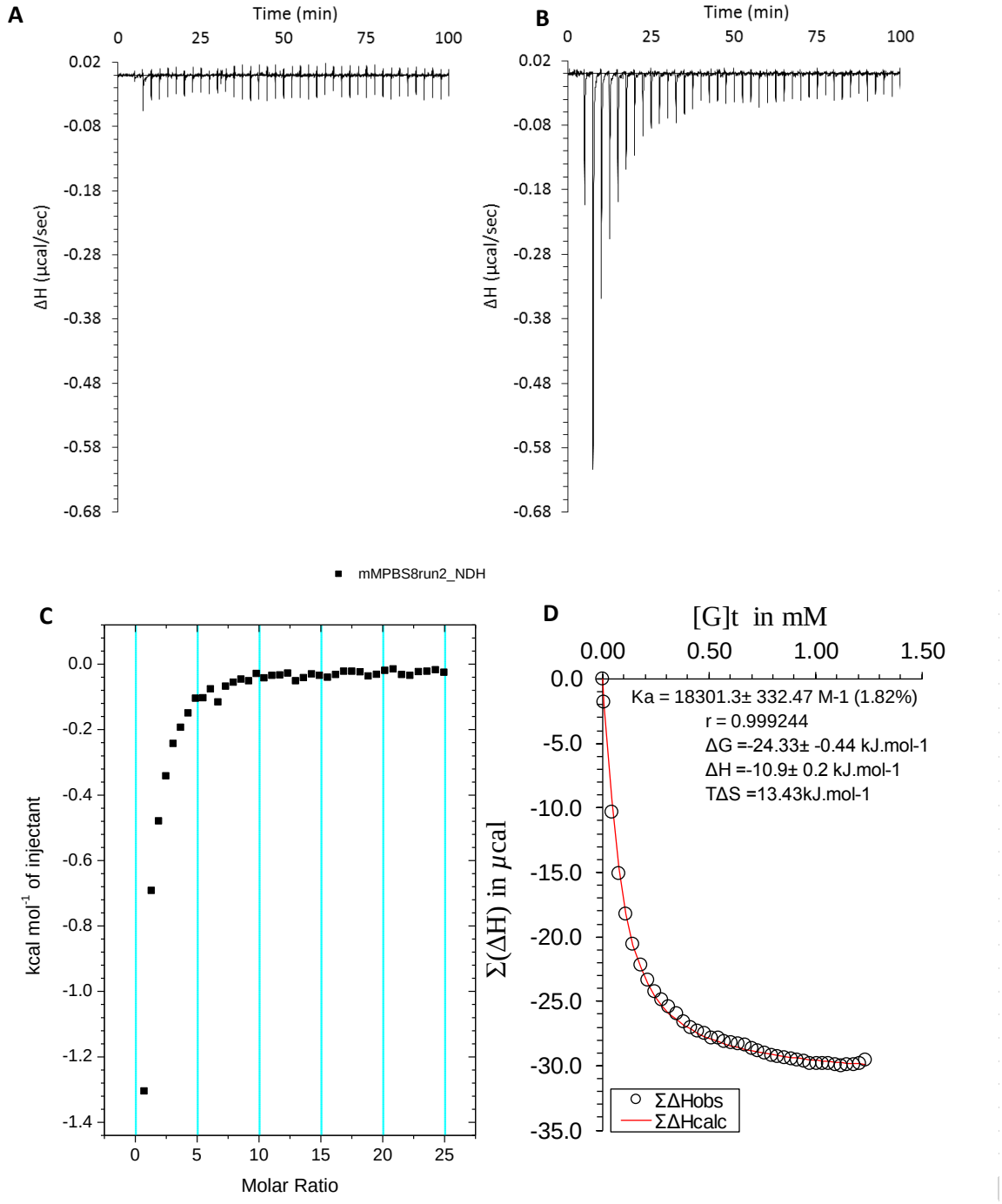


Figure 142 ITC binding results for receptor **3** (0.06 mM) titrated with D-glucose **4** (7 mM) in 10 mM phosphate buffered saline solution (pH 8), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 18,301 ± 332 M^{-1}).

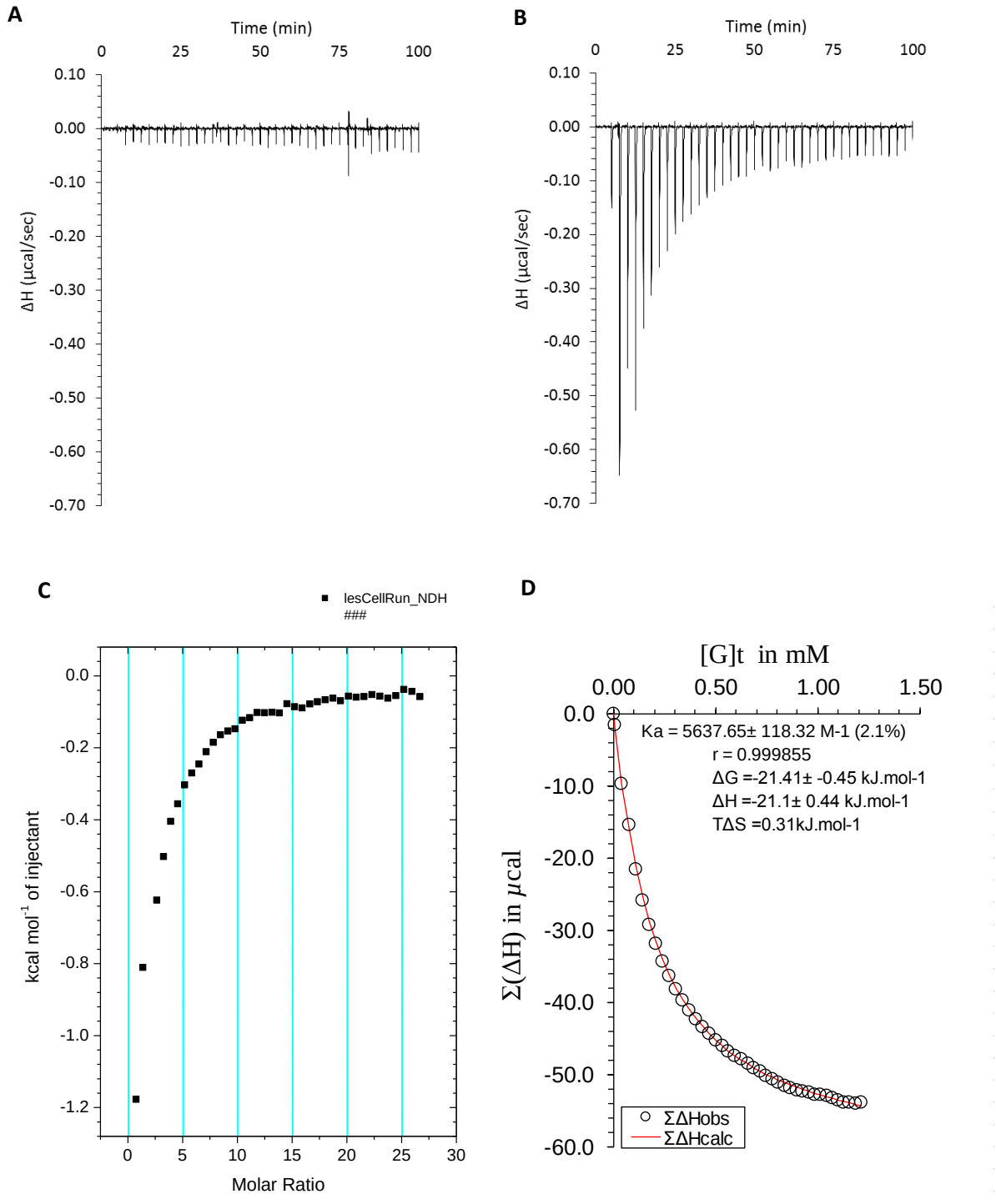


Figure 143 ITC binding results for receptor **3** (0.06 mM) titrated with D-glucose **4** (7 mM) in DMEM Cell Culture Medium (no glucose, 10k MWCO, 90% v/v) and 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into medium); B) shows the titration (sugar into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 5637 ± 118 M^{-1}).

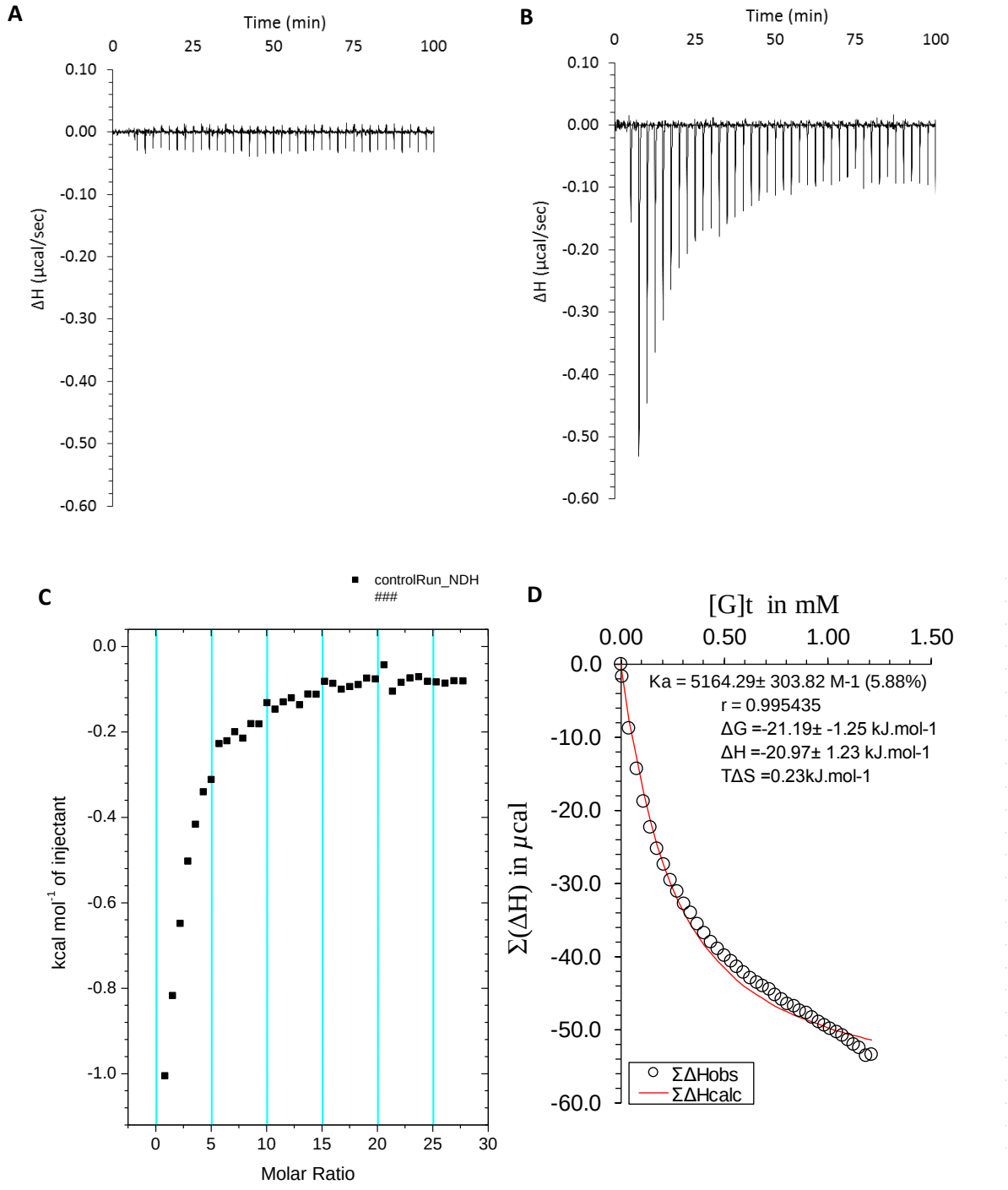


Figure 144 ITC binding results for receptor **3** (0.06 mM) titrated with D-glucose **4** (7 mM) in DMEM salt control medium with 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor 2); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 5164 ± 303 M^{-1}**).**

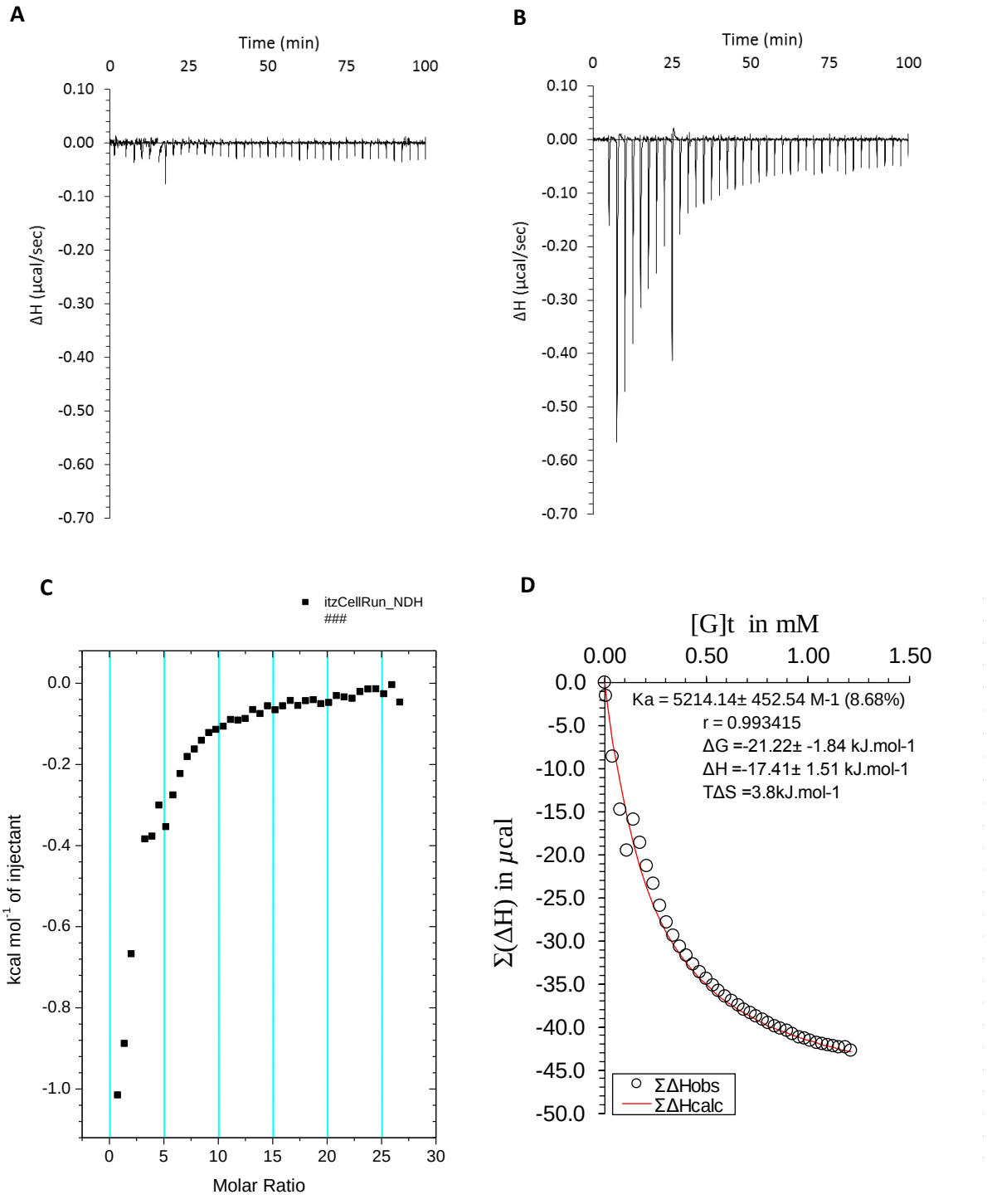


Figure 145 ITC binding results for receptor **3** (0.06 mM) titrated with D-glucose **4** (7 mM) in Leibovitz's L-15 Cell Culture Medium (no glucose, 10k MWCO, 90% v/v) and 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into medium); B) shows the titration (sugar into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 5214 ± 452 M^{-1}).

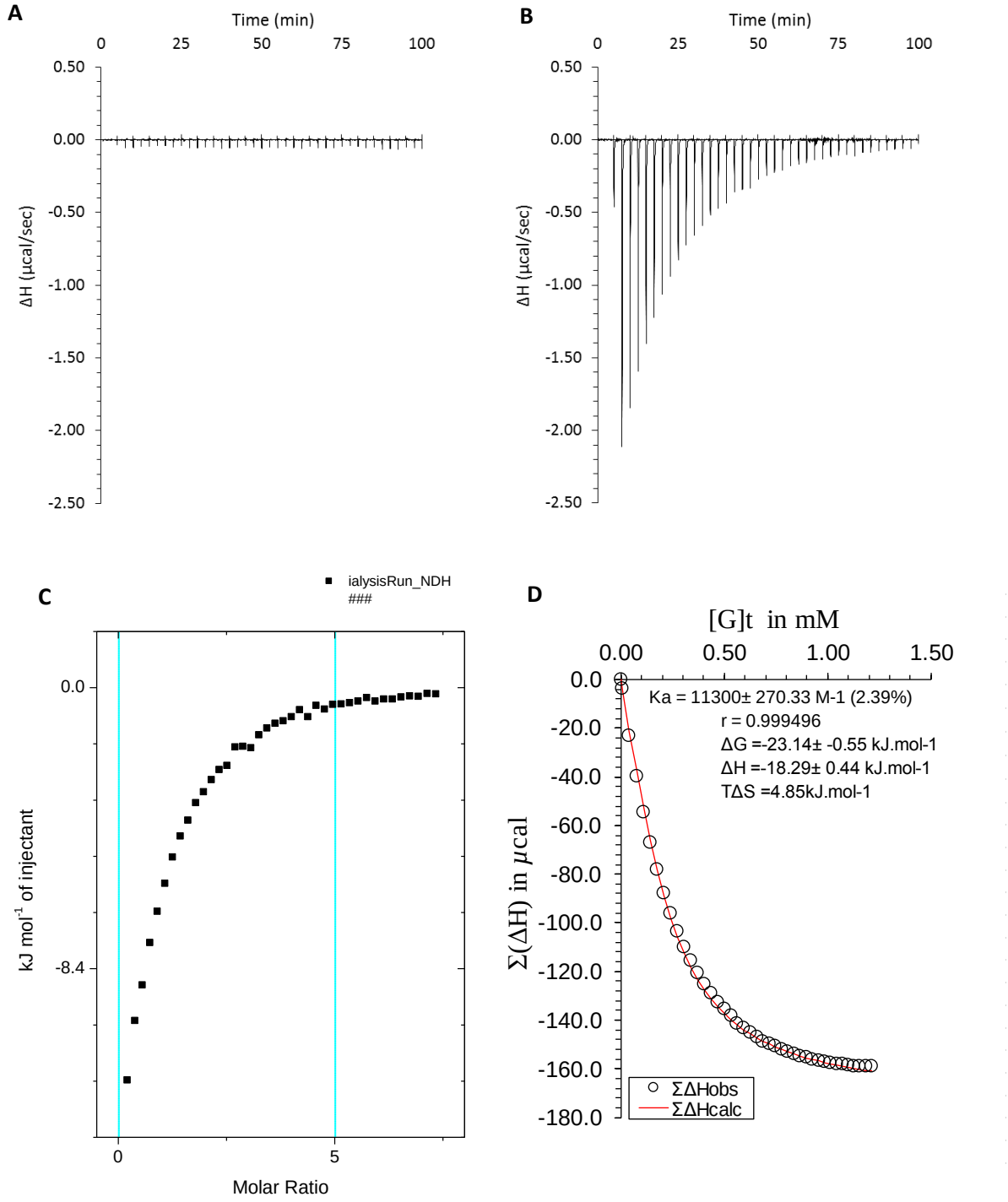


Figure 146 ITC binding results for receptor **3** (0.2 mM) titrated with D-glucose **4** (7 mM) in human serum (90% v/v, no glucose) with 10 mM phosphate buffer (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 11,300 ± 270 M^{-1}).

9.2.3 Binding studies for triethylbenzene receptor (3) with other susbtrates

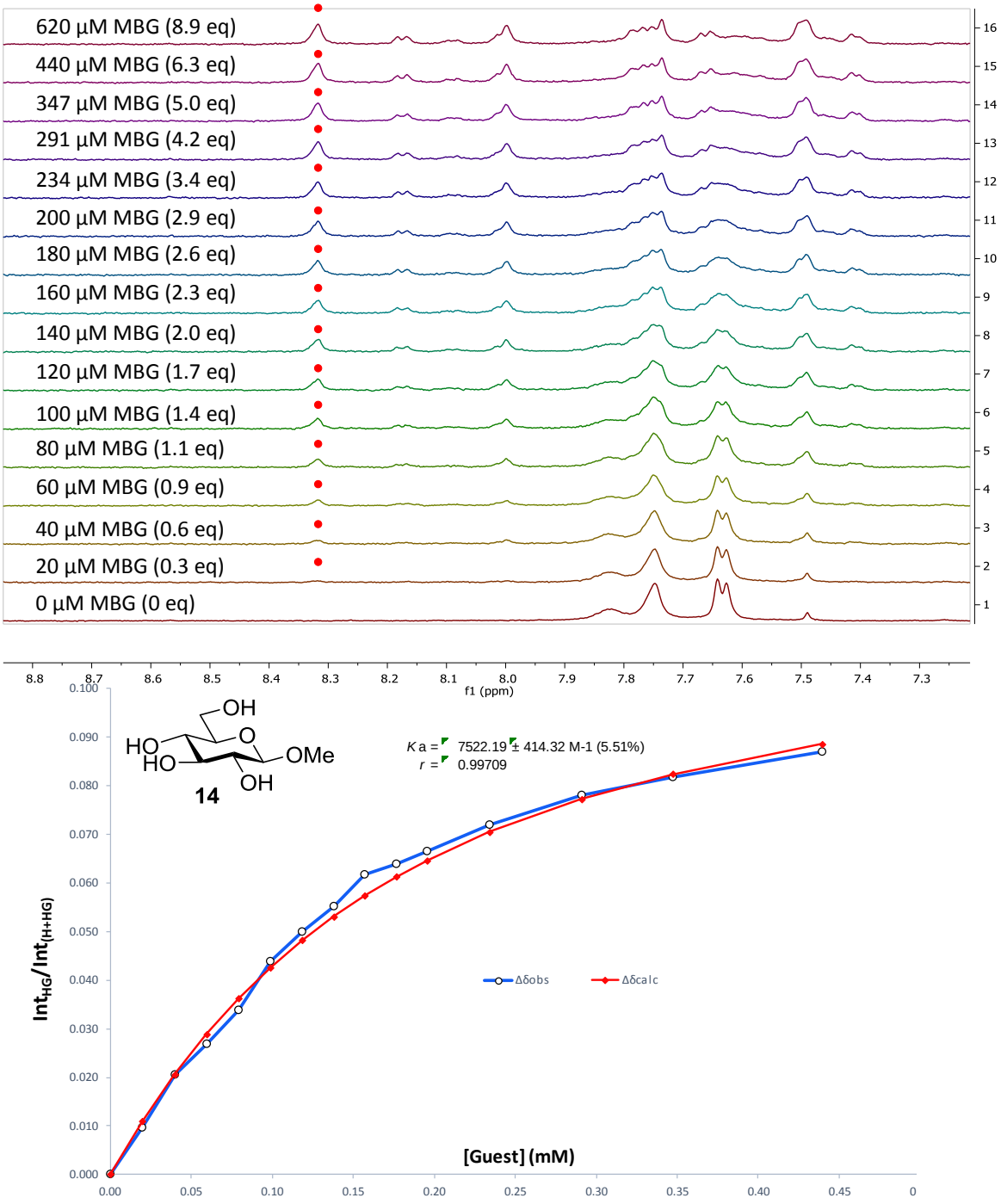


Figure 147 ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.07 mM) titrated with a combined solution of methyl β-D-glucoside **14** (10 mM) and receptor **3** (0.07 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with slow exchange on NMR timescale. Integrations of peak at 8.31 ppm (denoted with •) versus region 8.36-7.36 ppm were plotted against guest concentration (mM). The calculated values for the integrals are overlaid with the observed values, giving K_a = 7522 ± 414 M^{-1} (5.51%)

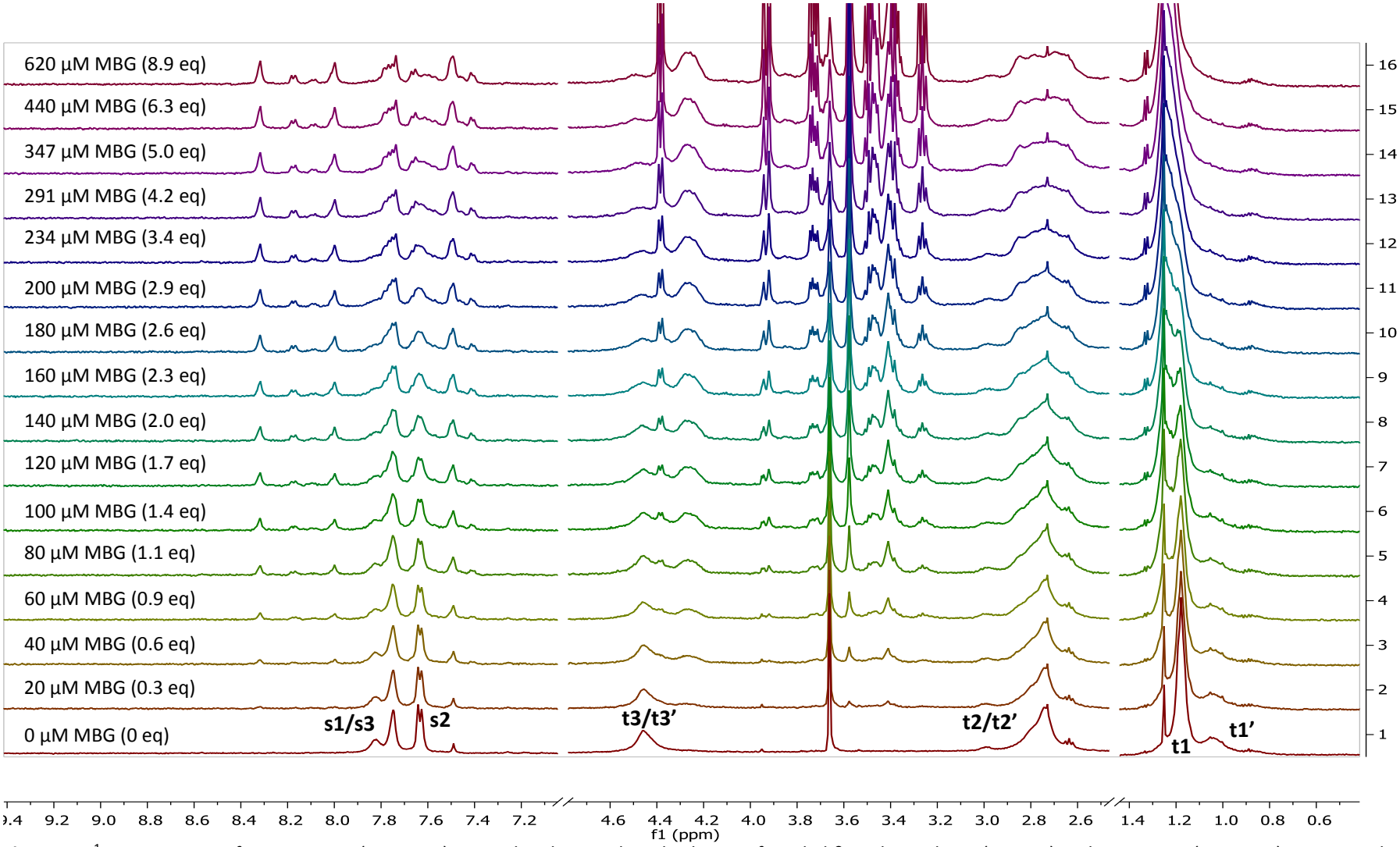


Figure 148[1]H NMR spectra for receptor **3** (0.07 mM) titrated with a combined solution of methyl β-D-glucoside **14** (10 mM) and receptor **3** (0.07 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

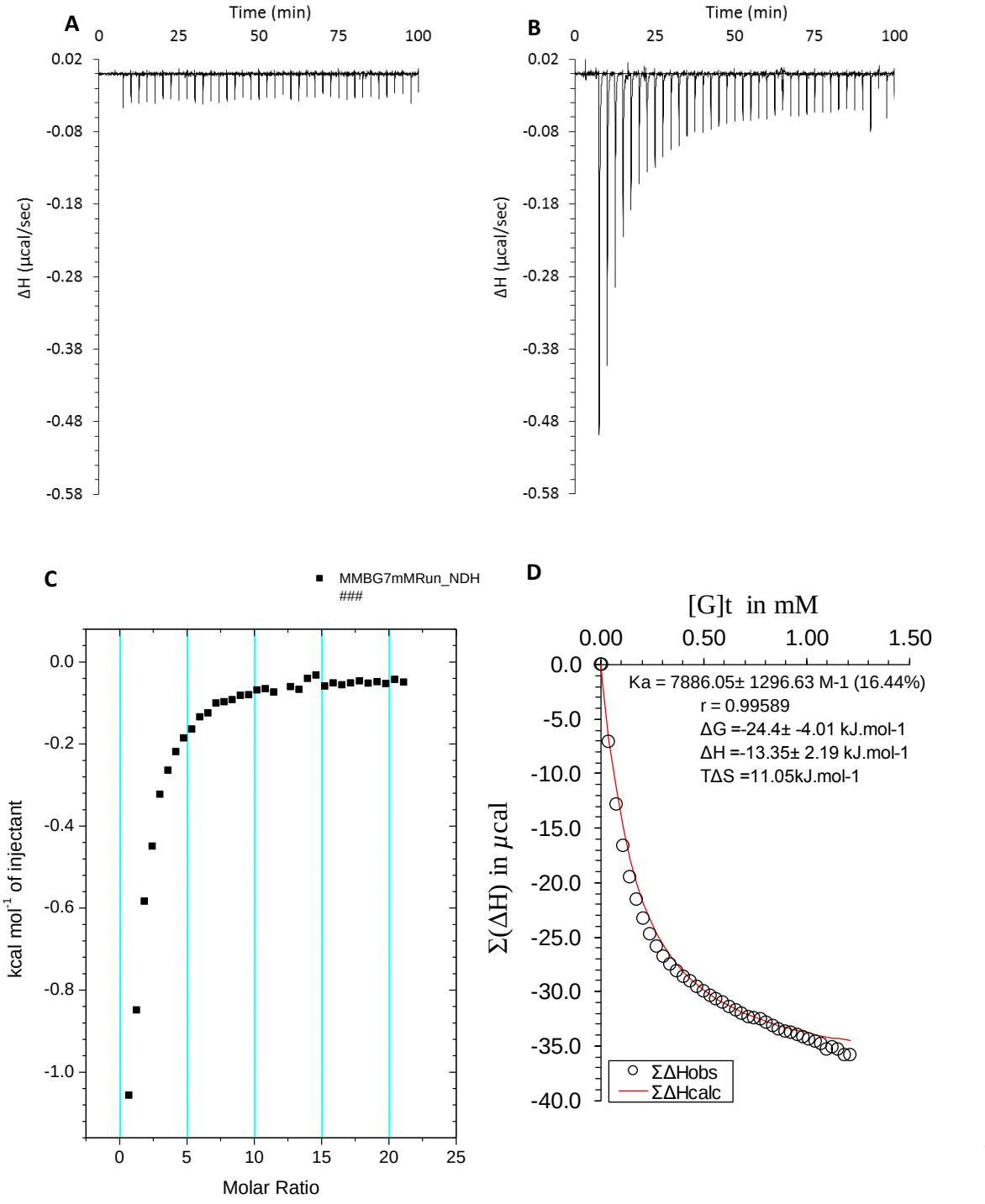


Figure 149ITC binding results for receptor **3** (0.06 mM) titrated with methyl β-D-glucoside **14** (7 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 7886 ± 1296 M^{-1}).

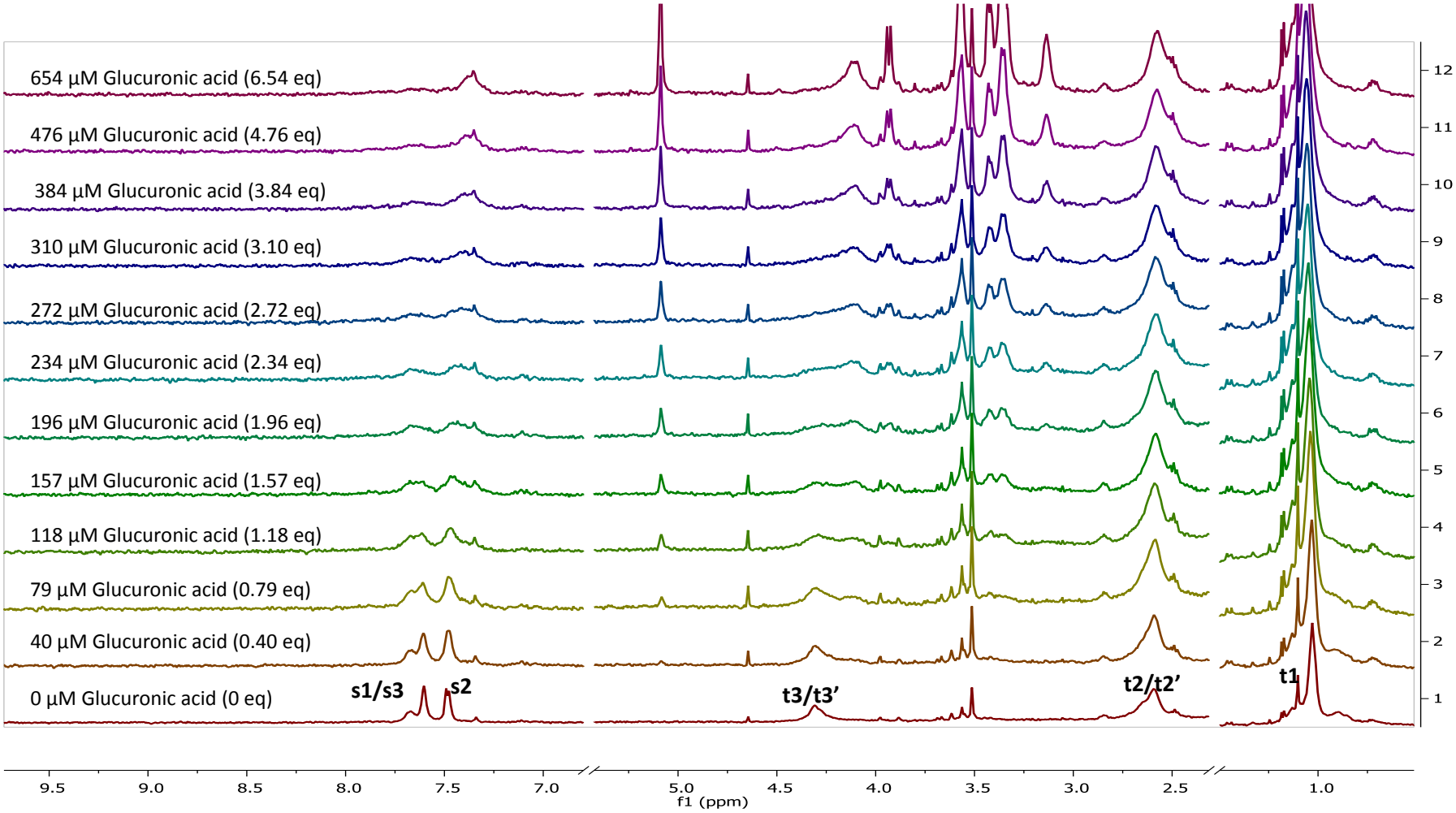


Figure 150 Partial ^{1}H NMR spectra for receptor **3** (0.1 mM) titrated with a combined solution of D-glucuronic acid **168** (10 mM) and receptor **3** (0.1 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with intermediate rate of exchange (rate between fast and slow exchange rates between H and HG species) on NMR timescale. Due to severe broadening of peaks for receptor **3** upon addition of guest, no K_a was determinable.

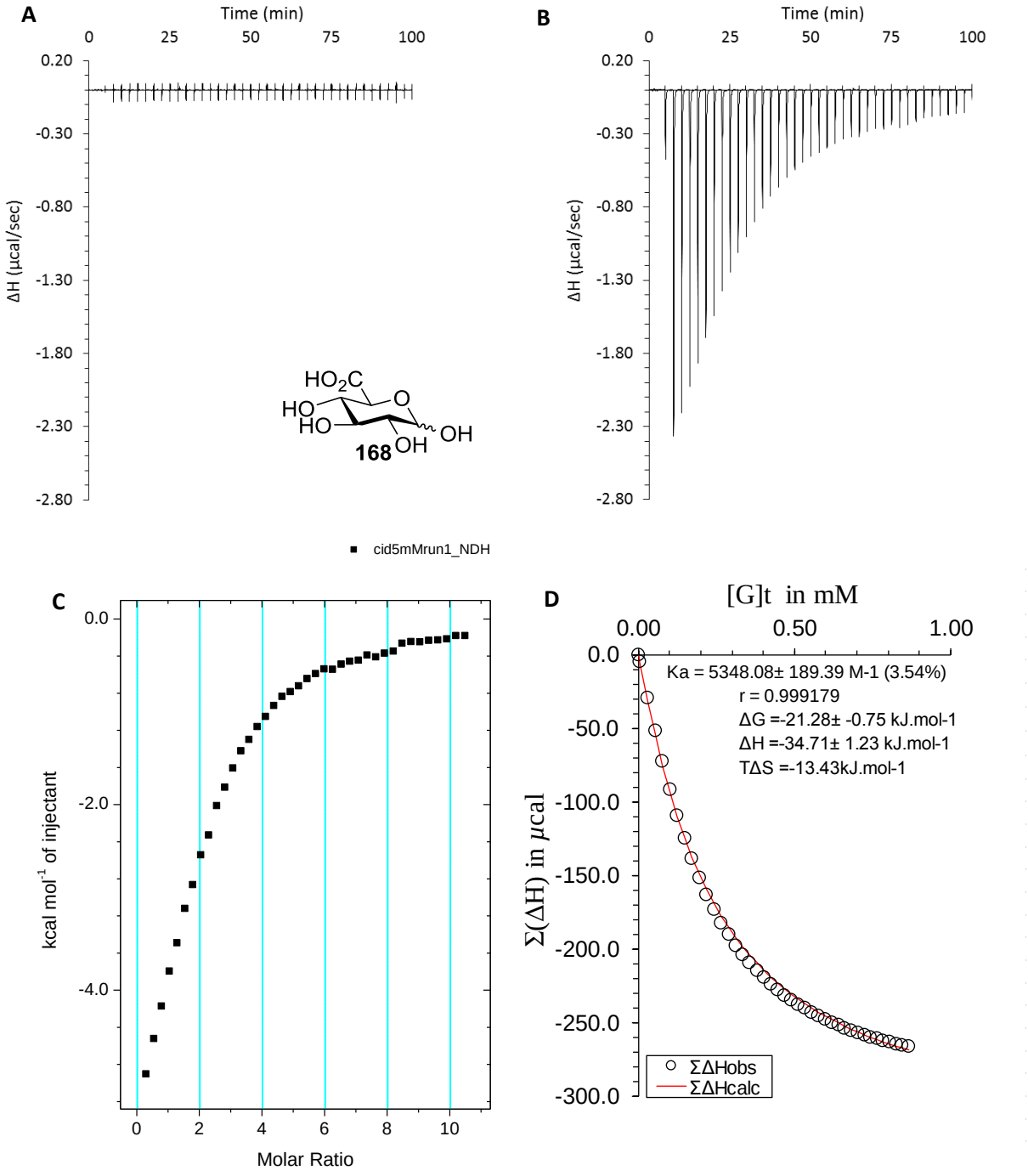


Figure 151 ITC binding results for receptor **3** (0.2 mM) titrated with D-glucuronic acid **168** (5 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 5348 ± 189 M^{-1}).

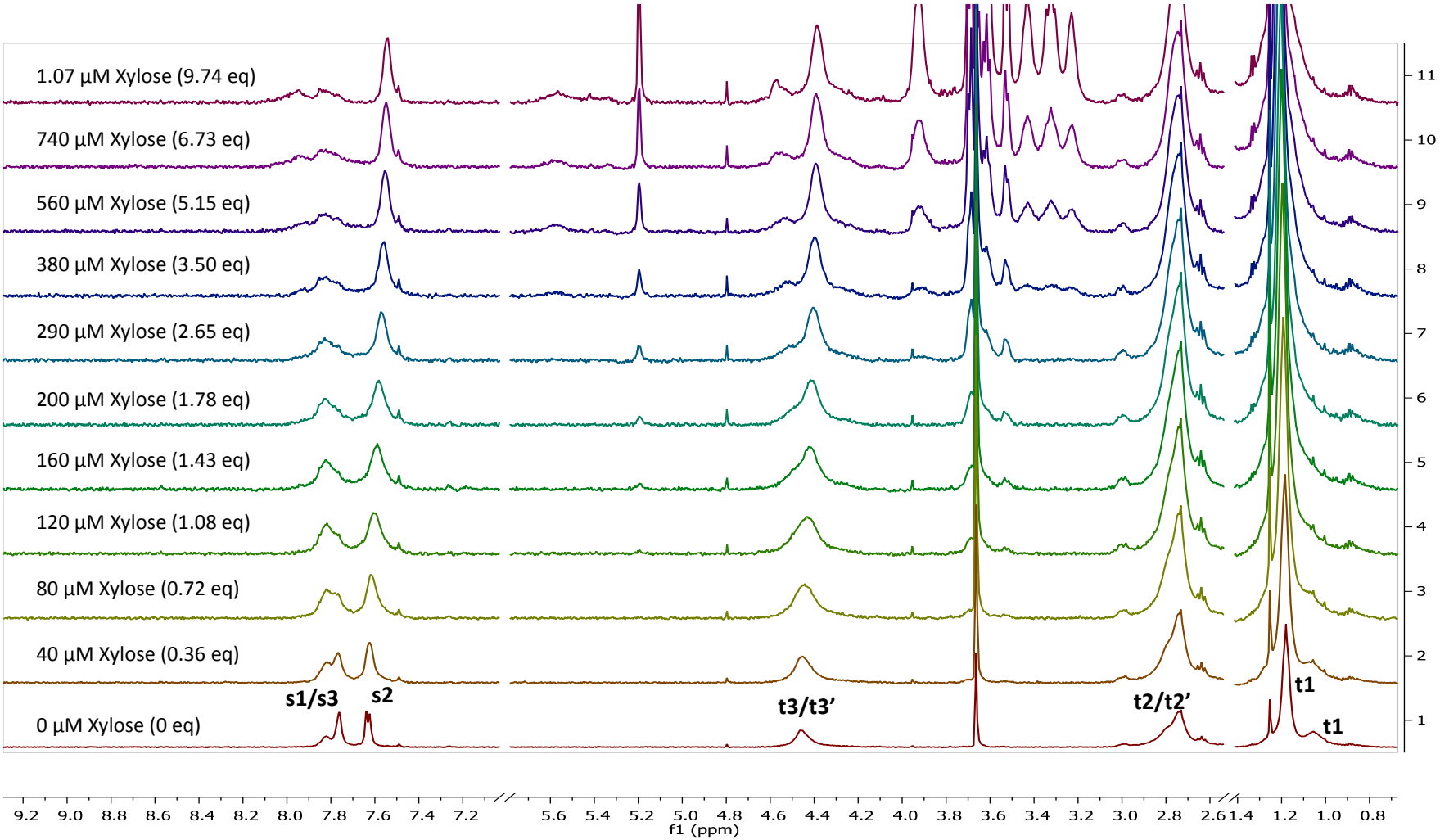


Figure 152 ^{1}H NMR spectra for receptor **3** (0.11 mM) titrated with a combined solution of D-xylose **149** (10 mM) and receptor **3** (0.11 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) a t 298 K. Spectra imply binding with intermediate rate of exchange (rate between fast and slow exchange rates between H and HG species) on NMR timescale. Due to severe broadening of peaks for receptor **3** upon addition of guest, no K_a was determinable.

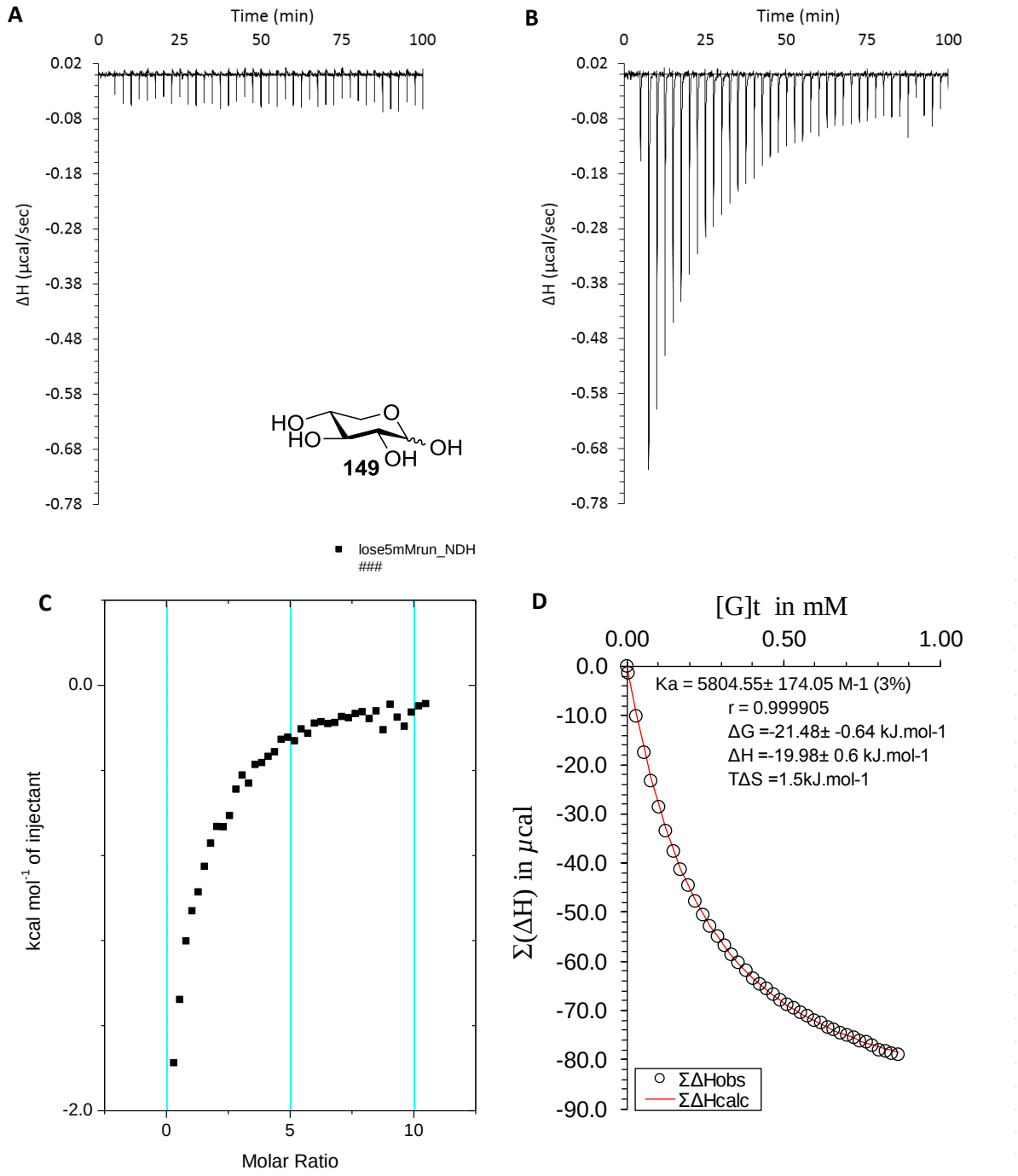


Figure 153 ITC binding results for receptor **3** (0.1 mM) titrated with D-xylose **149** (5 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 5804 ± 174 M^{-1}).

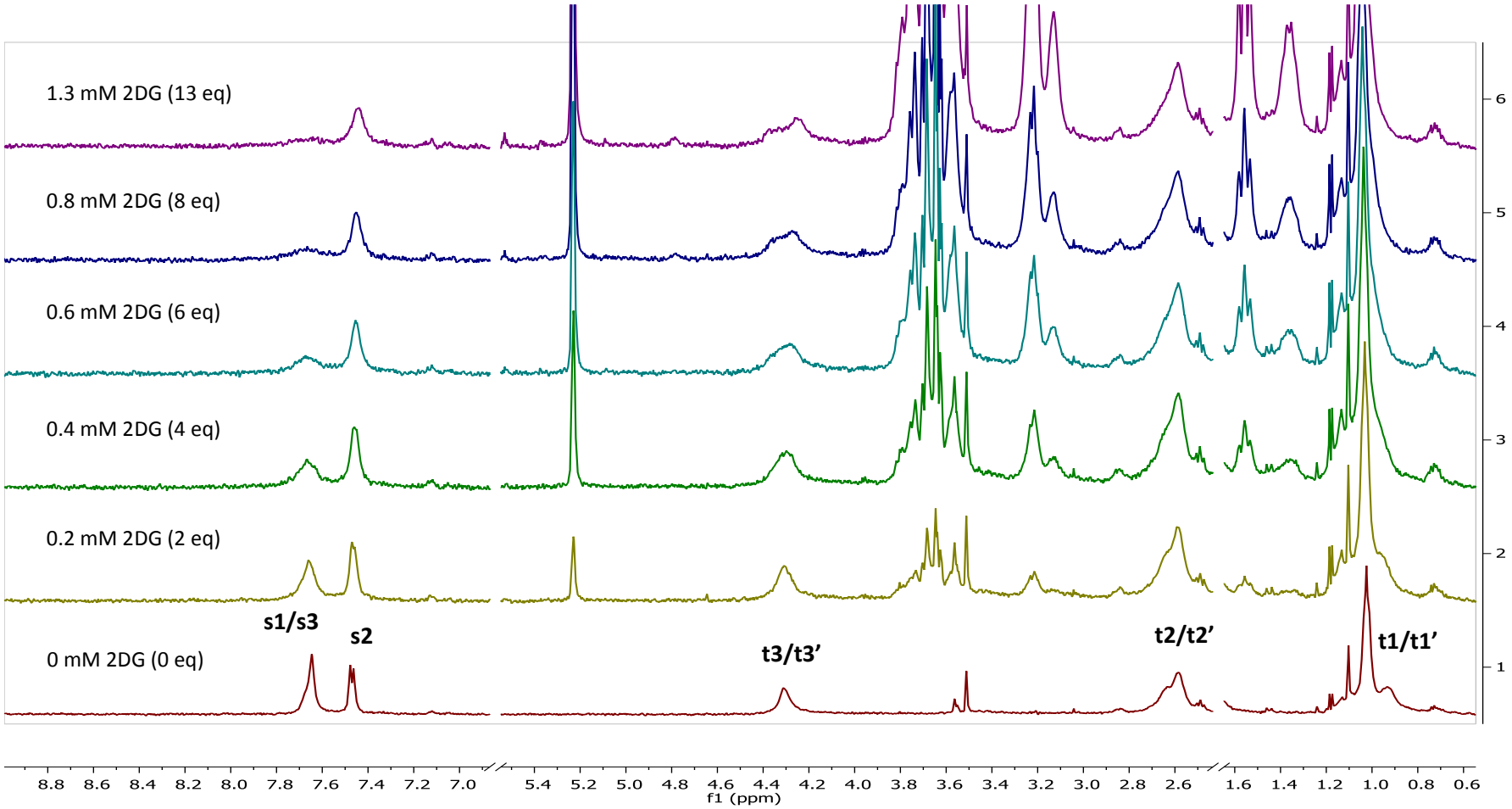


Figure 154 Partial ^{1}H NMR spectra for receptor **3** (0.1 mM) titrated with a combined solution of 2-deoxy-D-glucose **169** (50 mM) and receptor **3** (0.1 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with intermediate rate of exchange (rate between fast and slow exchange rates between H and HG species) on NMR timescale. Due to severe broadening of peaks for receptor **3** upon addition of guest, no K_a was determinable.

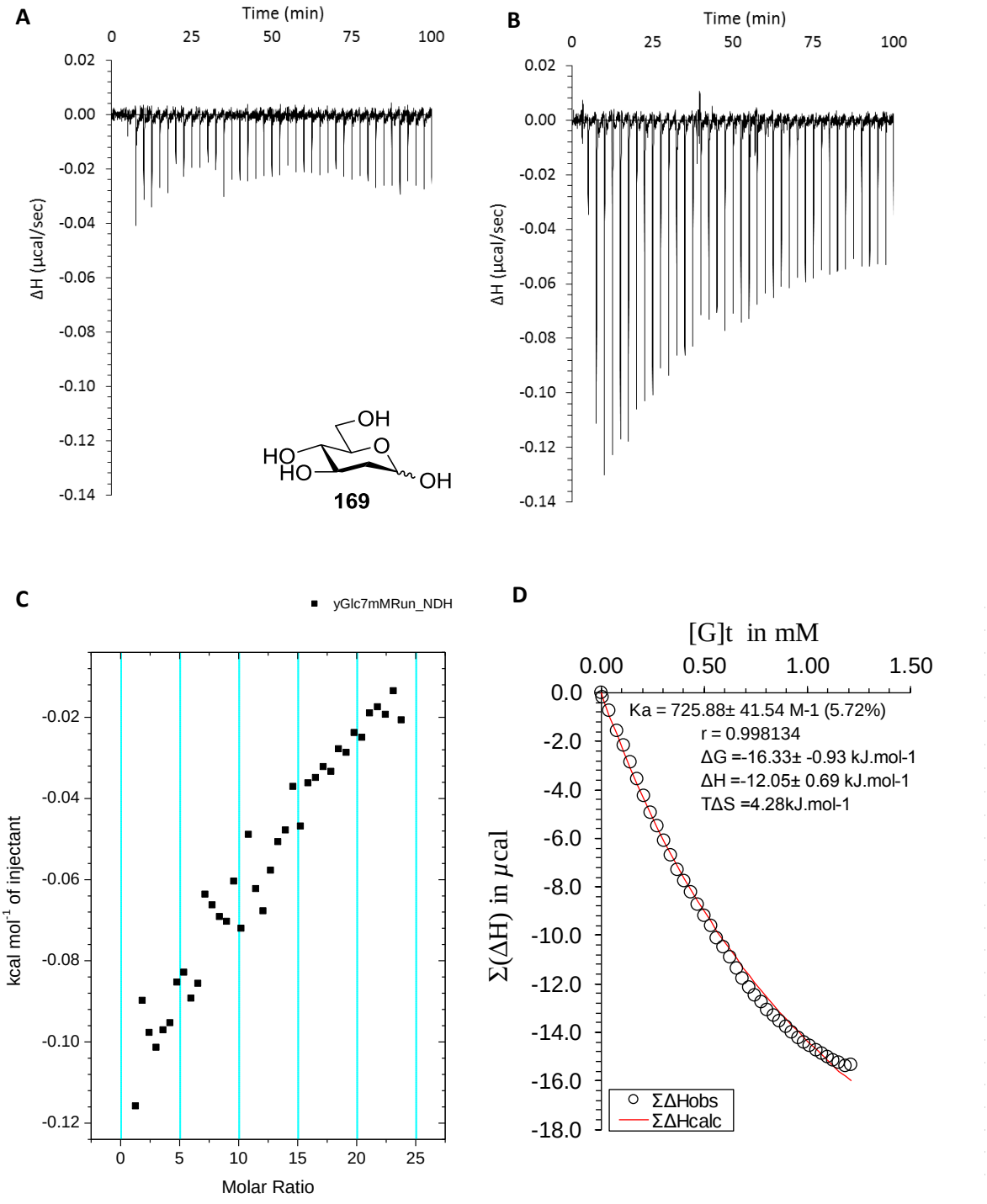


Figure 155 ITC binding results for receptor **3** (0.06 mM) titrated with 2-deoxy-D-glucose **169** (7 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 725 ± 41 M^{-1}).

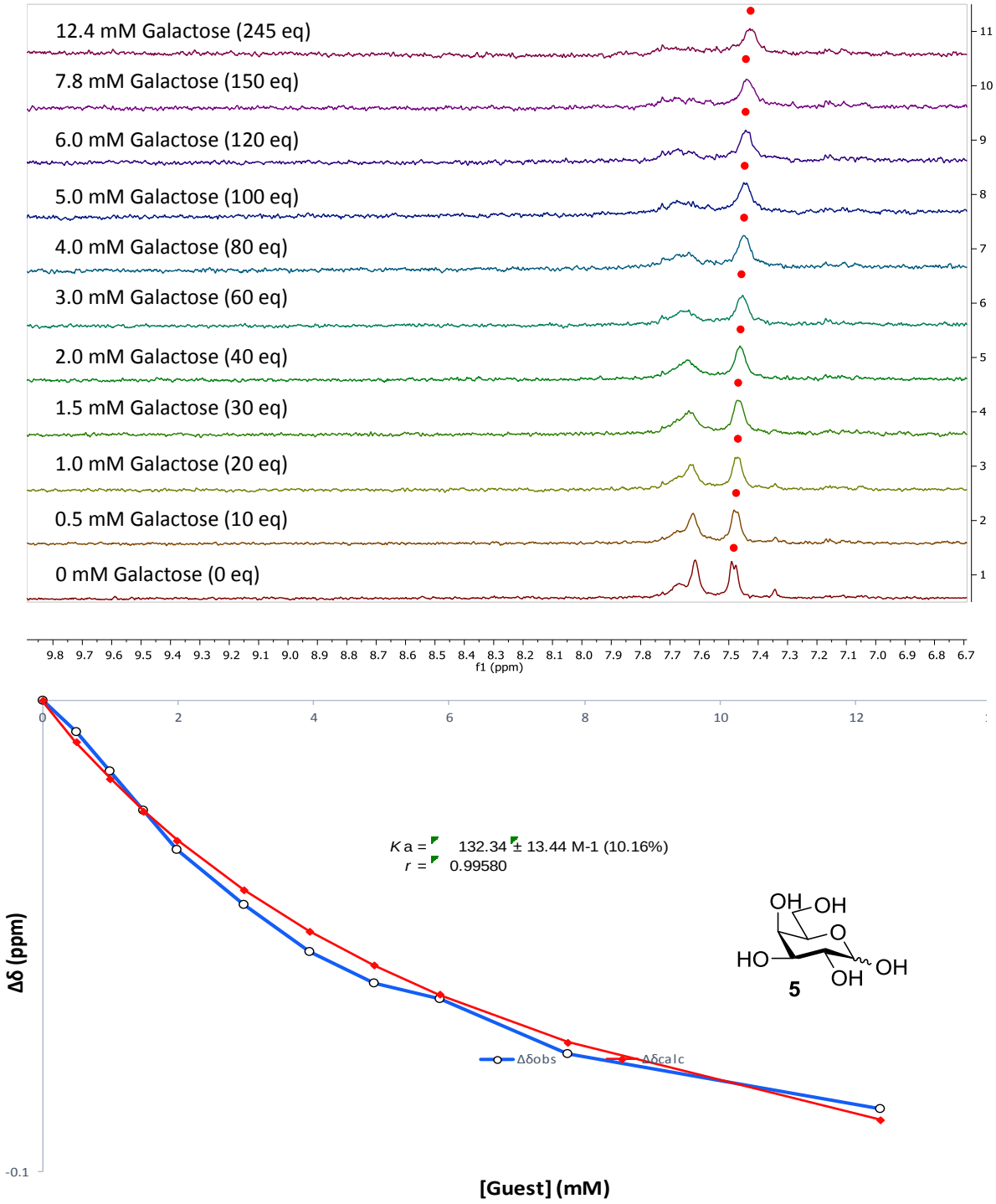


Figure 156 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.05 mM) titrated with a combined solution of D-galactose **5** (250 mM) and receptor 2 (0.05 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with fast/intermediate exchange on NMR timescale. Changes in chemical shift (Δδ ppm) of peak at 7.63ppm (denoted with •) were plotted against increasing guest concentration (mM). The calculated values for the Δδ are overlaid with the observed values giving K_a = 132 ± 13 M^{-1} (10.2%).

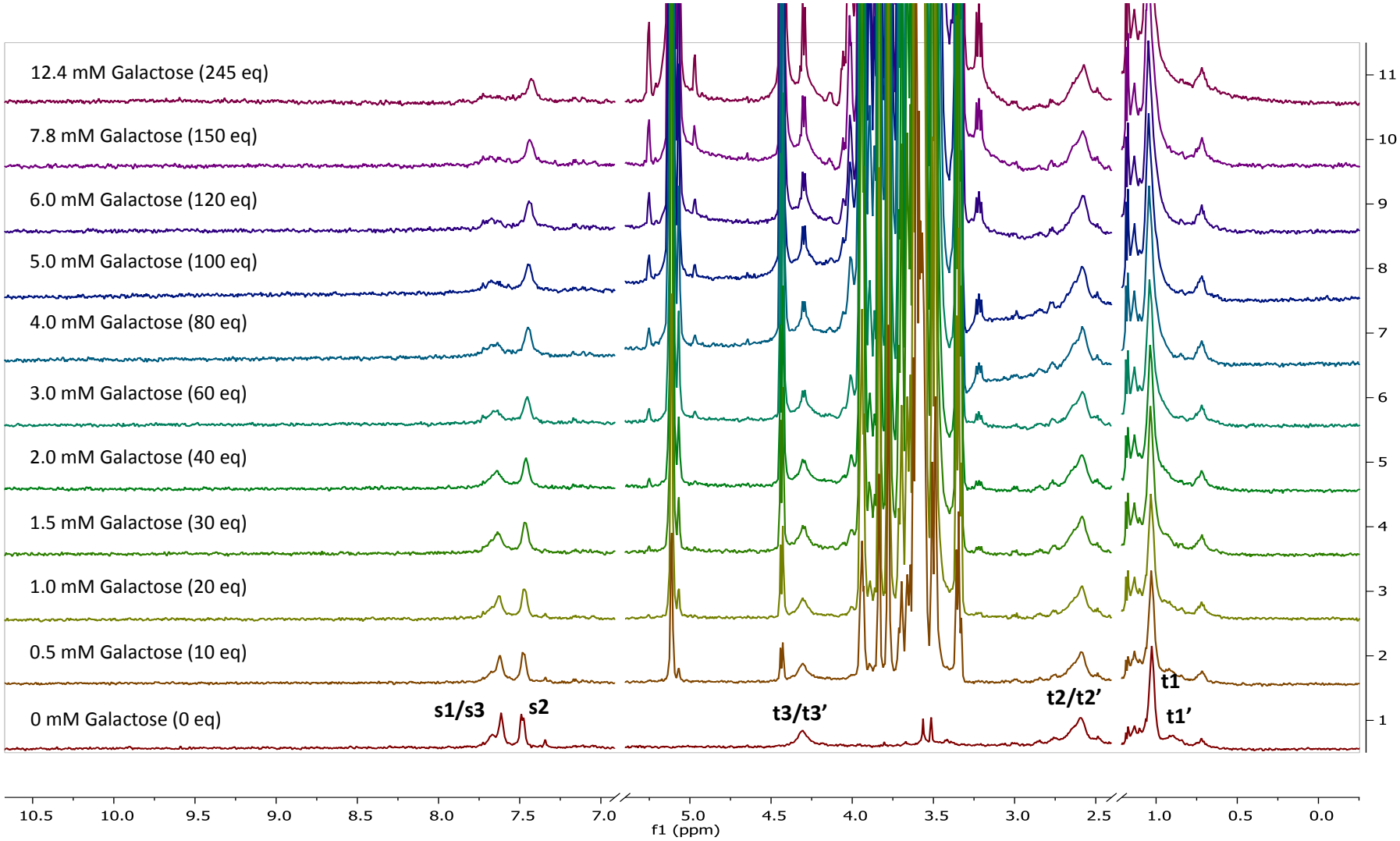


Figure 157 ^{1}H NMR spectra for receptor **3** (0.05 mM) titrated with a combined solution of D-galactose **5** (250 mM) and receptor **3** (0.05 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

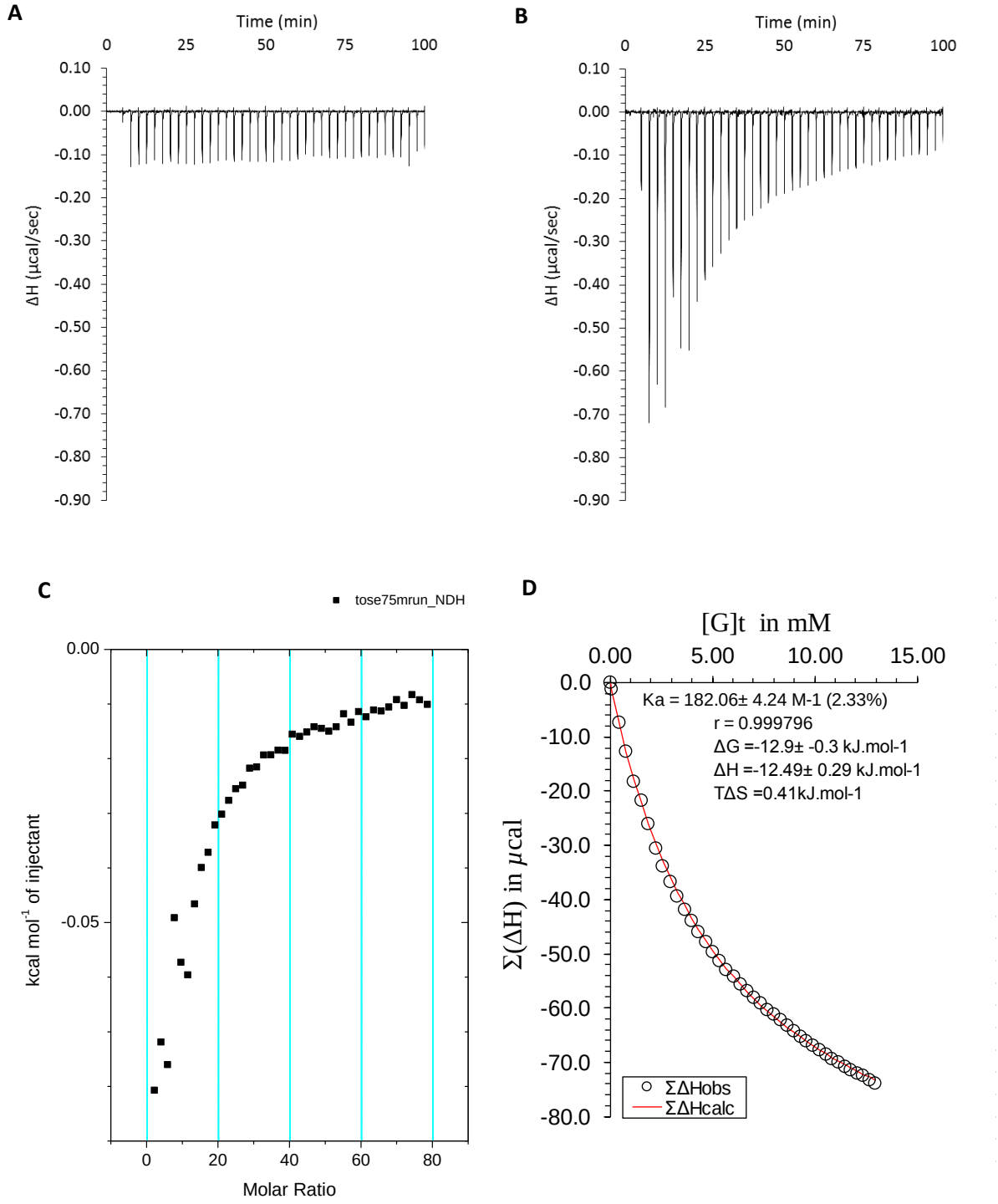


Figure 158 ITC binding results for receptor **3** (0.2 mM) titrated with D-galactose **5** (75 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 182 ± 4.2 M^{-1}).

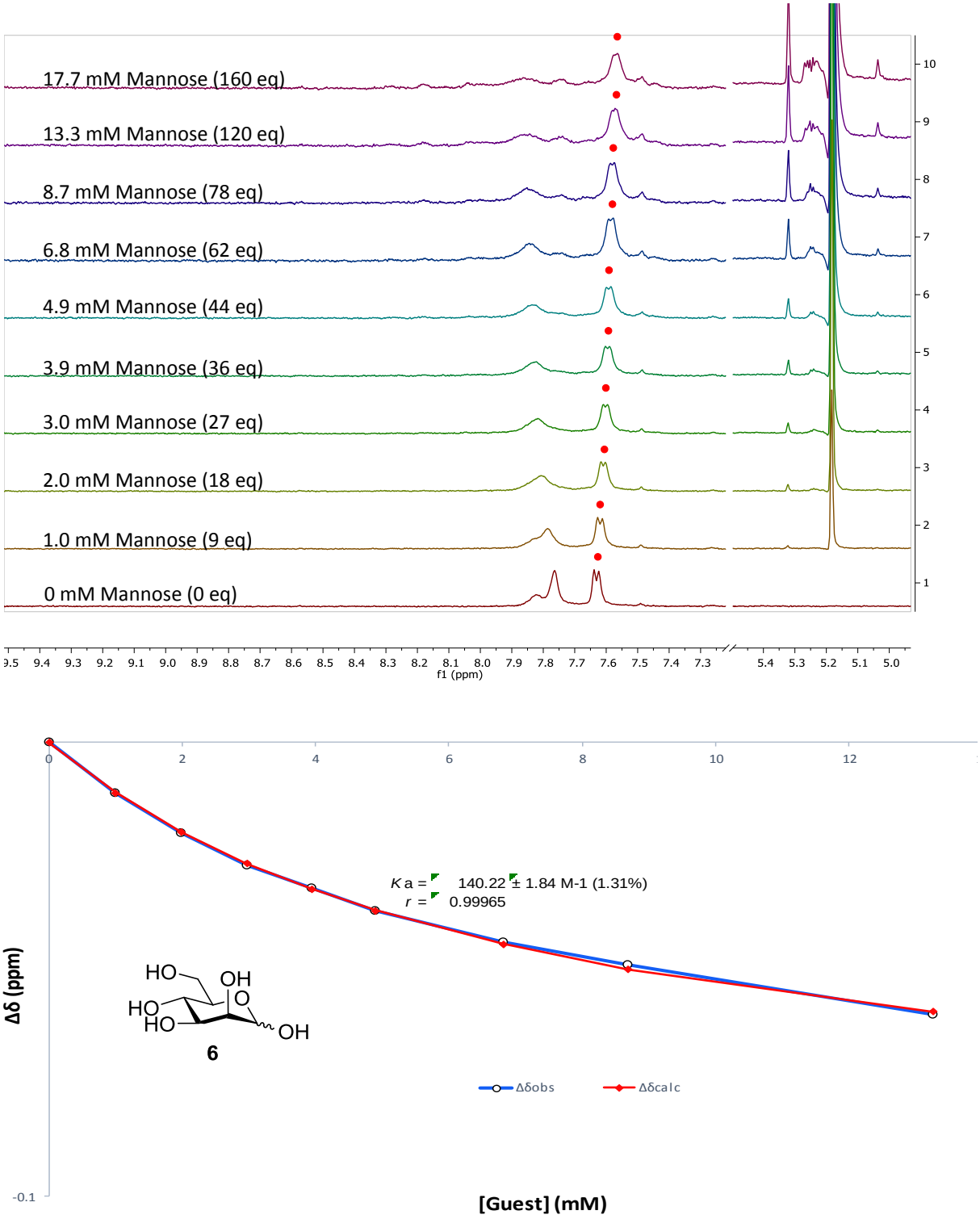


Figure 159 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.11 mM) titrated with a combined solution of D-mannose **6** (250 mM) and receptor 2 (0.11 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with fast/intermediate exchange on NMR timescale. Changes in chemical shift (Δδ ppm) of peak at 7.63 ppm (denoted with •) were plotted against increasing guest concentration (mM). The calculated values for the Δδ are overlaid with the observed values giving K_a = 140 ± 2 M^{-1} (1.31%)

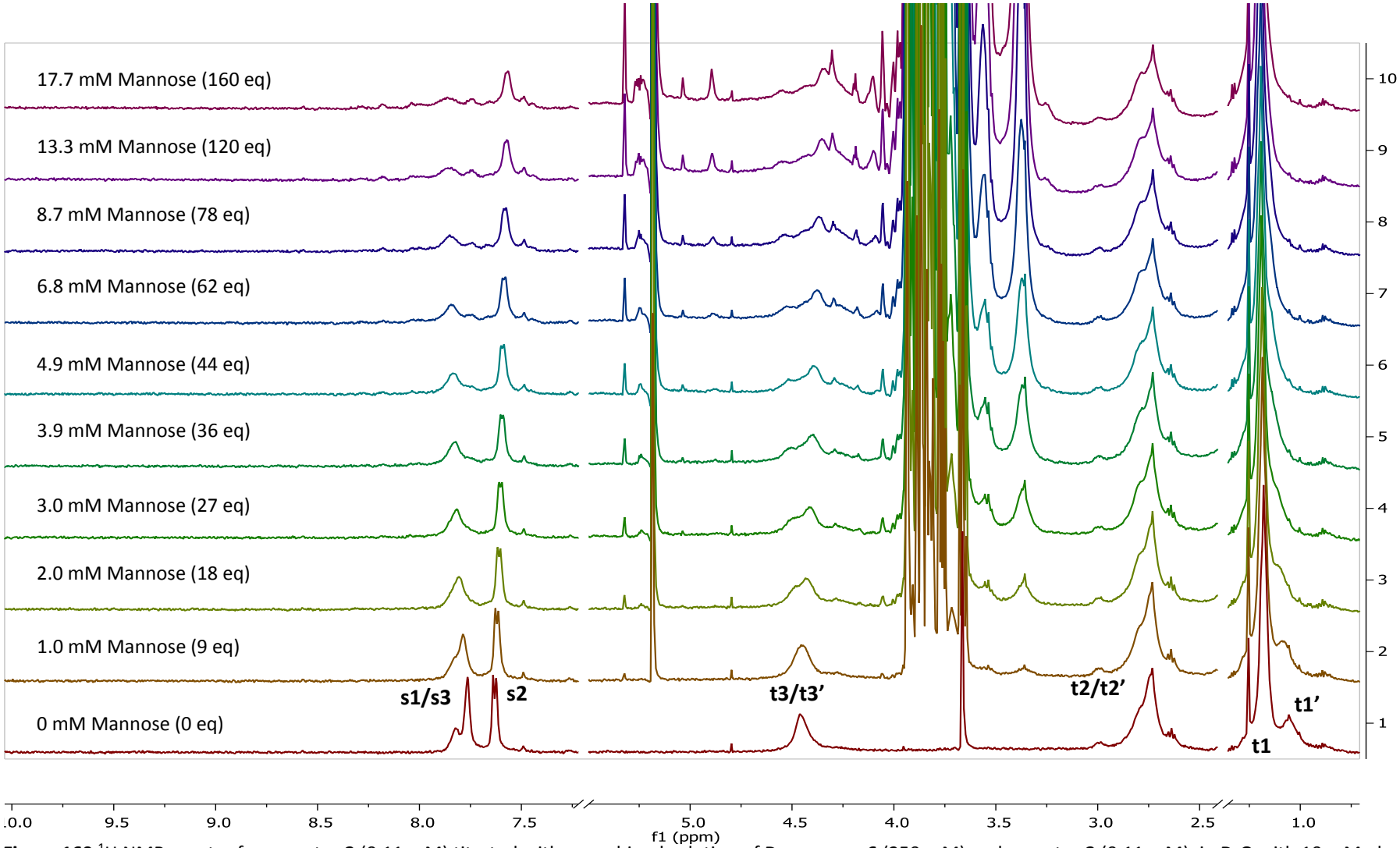


Figure 160 ^{1}H NMR spectra for receptor **3** (0.11 mM) titrated with a combined solution of D-mannose **6** (250 mM) and receptor **3** (0.11 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

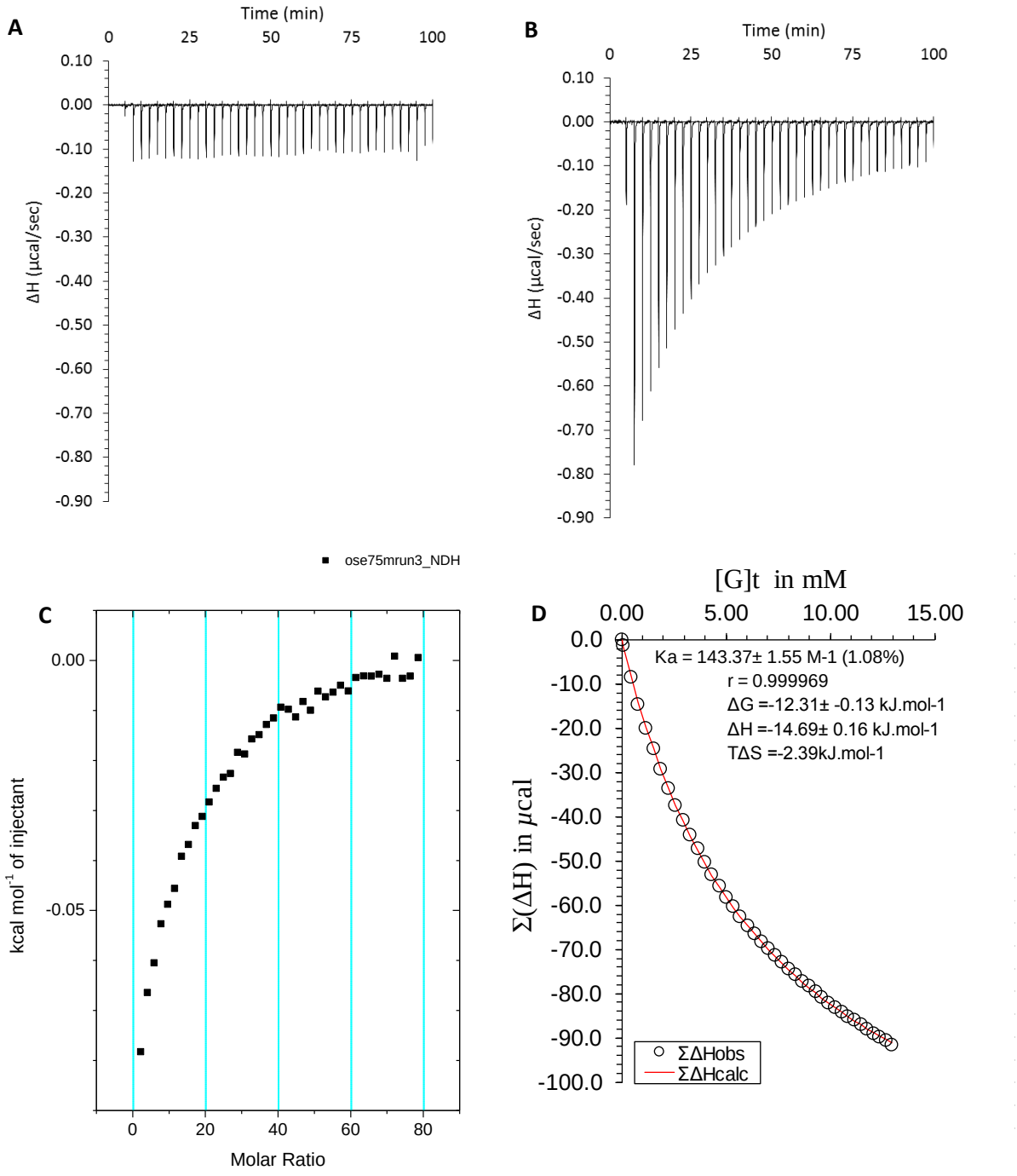


Figure 161 ITC binding results for receptor **3** (0.2 mM) titrated with D-mannose **6** (75 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 143 ± 1.5 M^{-1})

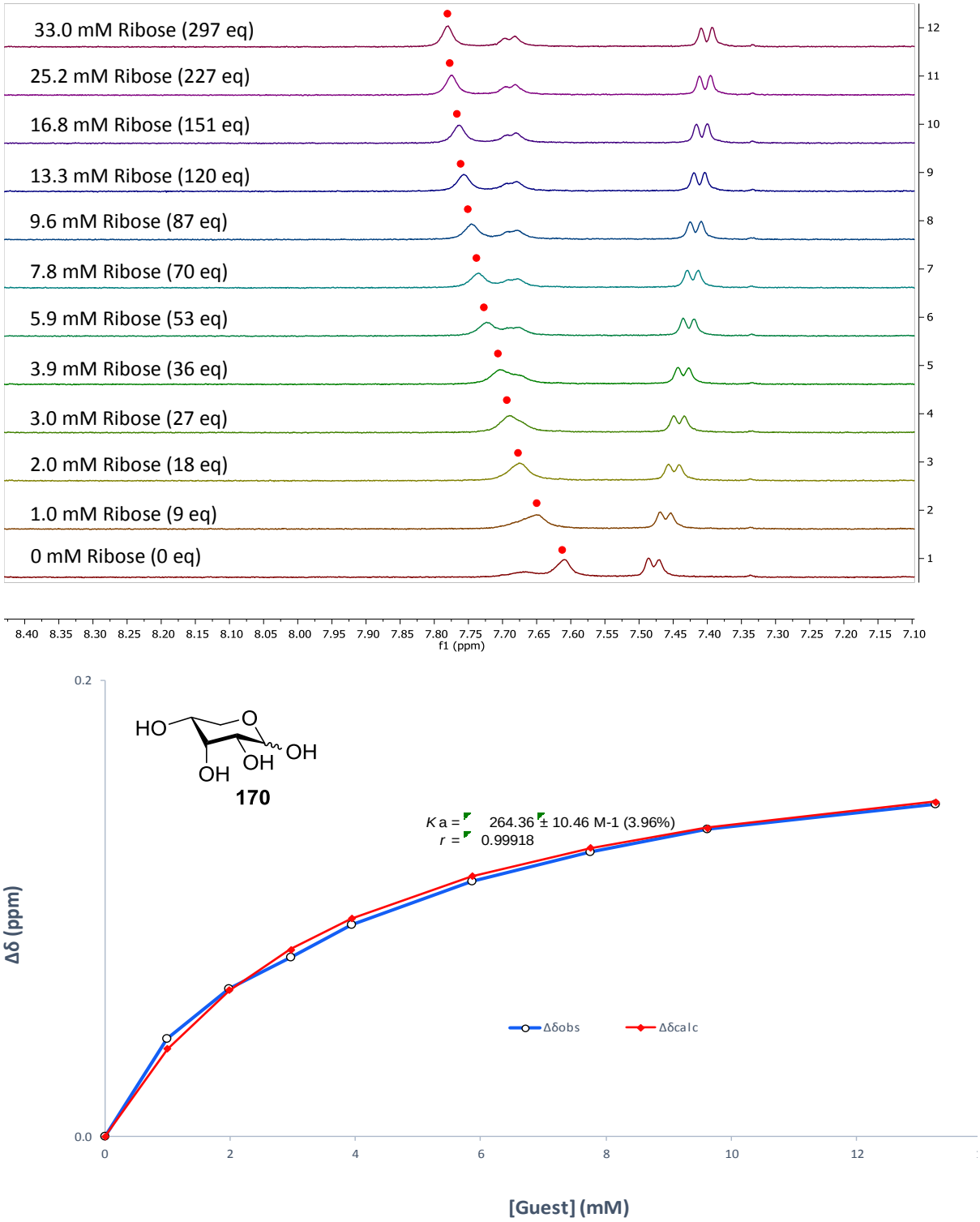


Figure 162 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.11 mM) titrated with a combined solution of D-ribose **170** (250 mM) and receptor **3** (0.11 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with fast exchange on NMR timescale. Changes in chemical shift (Δδ ppm) of peak at 7.83ppm (denoted with •) were plotted against increasing guest concentration (mM). The calculated values for the Δδ are overlaid with the observed values giving K_a = 264 ± 10 M^{-1} (3.96%).

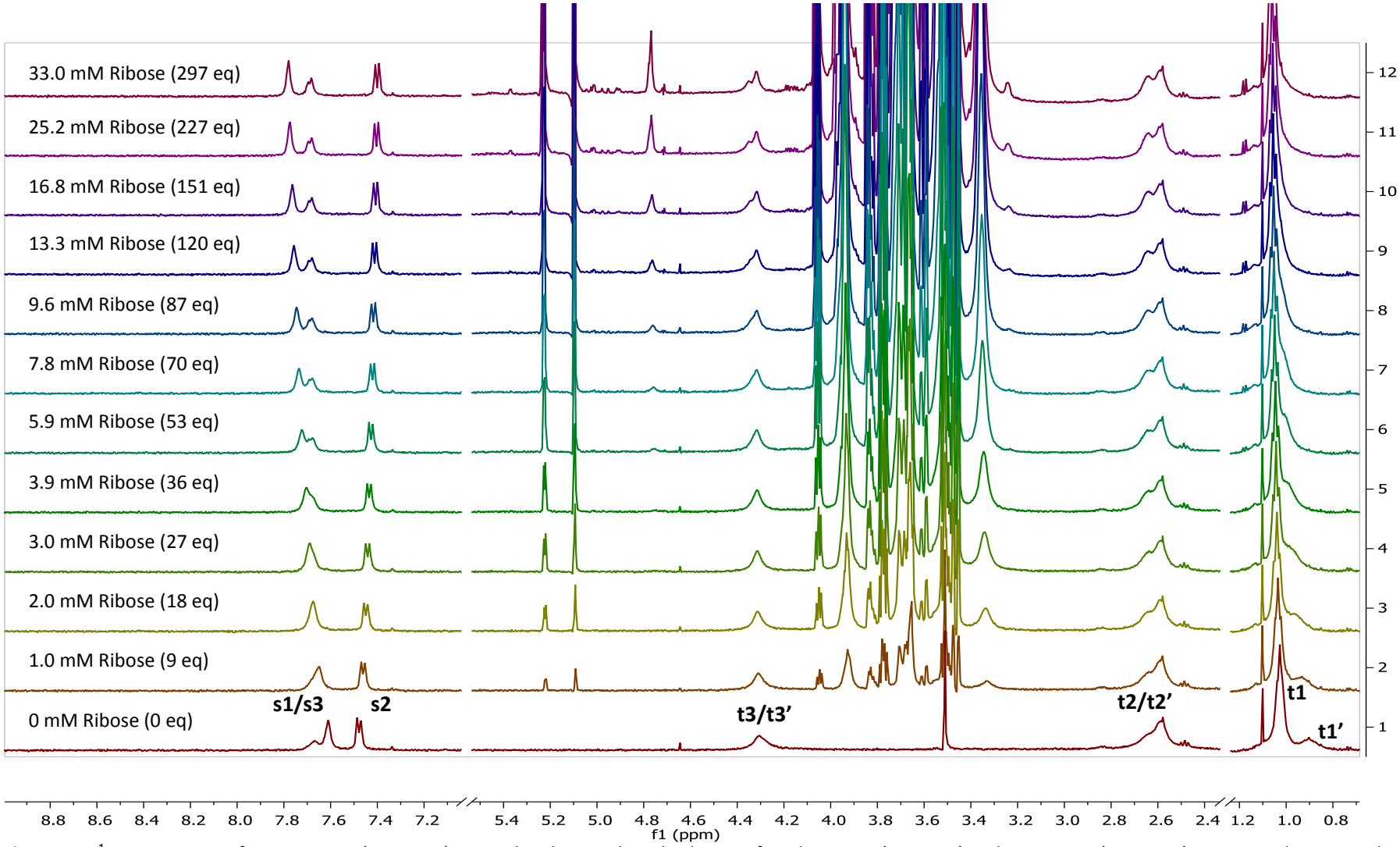


Figure 163 ^{1}H NMR spectra for receptor **3** (0.11 mM) titrated with a combined solution of D-ribose **170** (250 mM) and receptor **3** (0.11 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

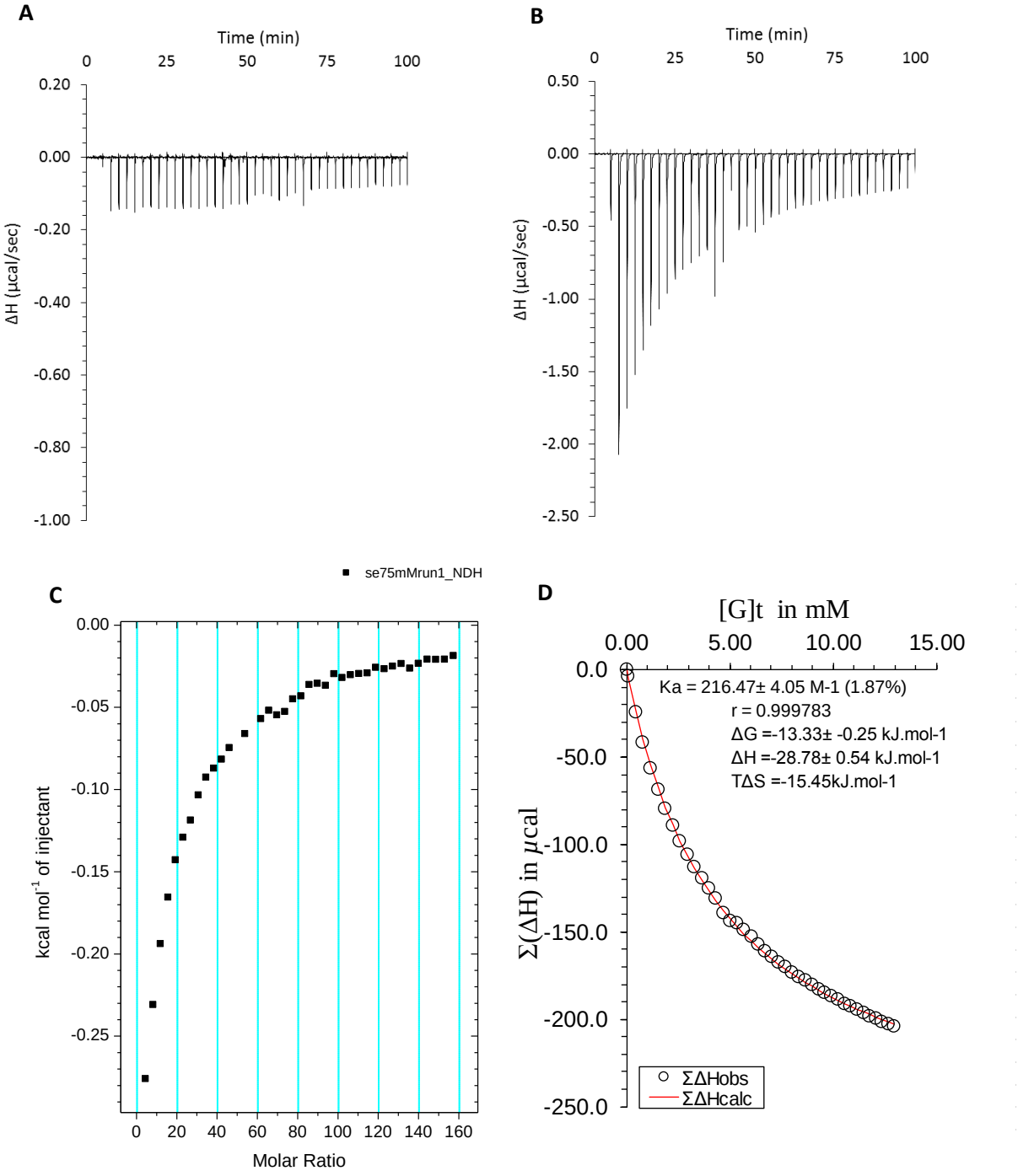


Figure 164 ITC binding results for receptor **3** (0.1 mM) titrated with D-ribose **170** (75 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 216.5 ± 4.1 M^{-1}).

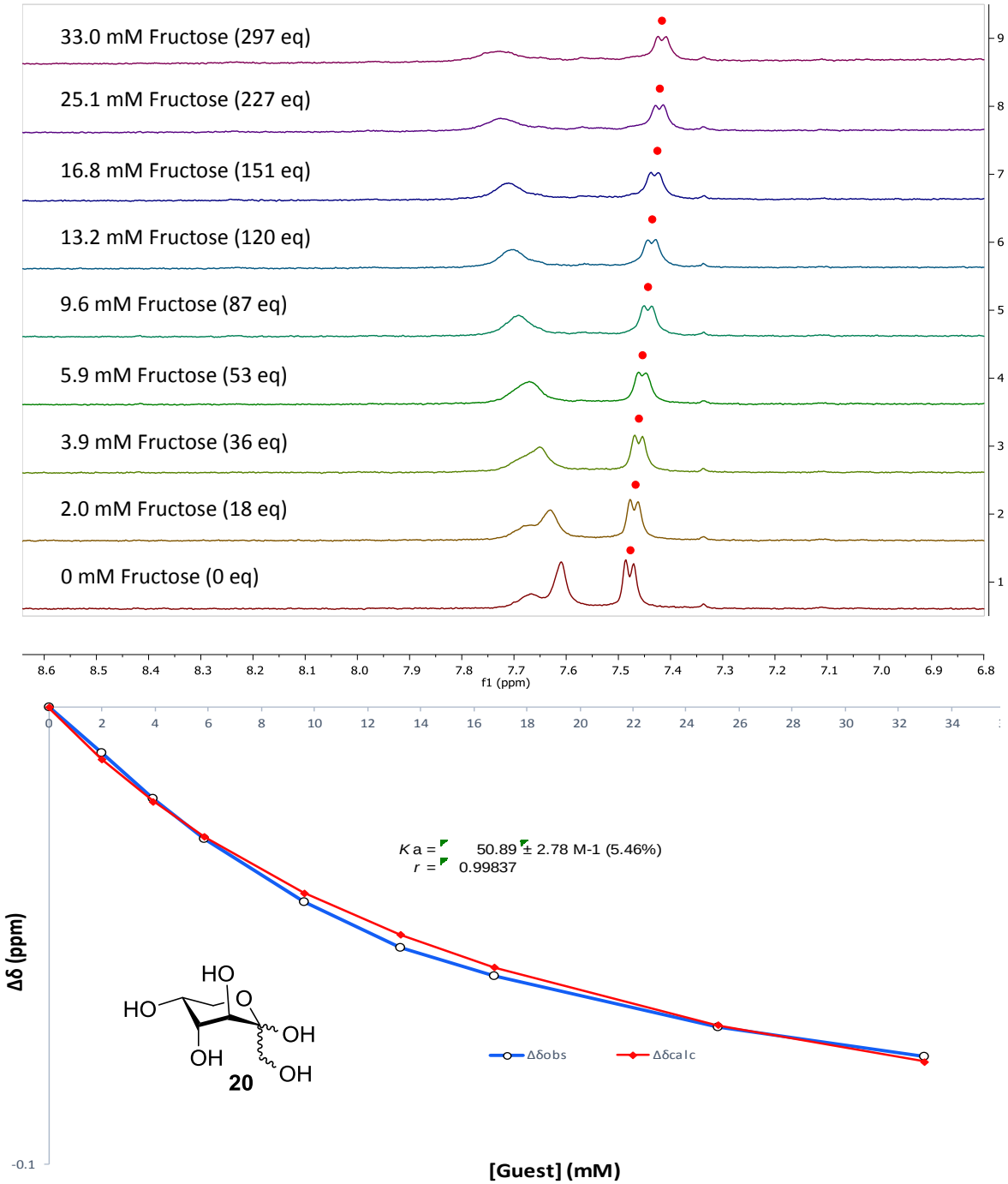


Figure 165 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.11 mM) titrated with a combined solution of D-fructose **20** (250 mM) and receptor 2 (0.11 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with fast/intermediate exchange on NMR timescale. Changes in chemical shift (Δδ ppm) of peak at 7.63ppm (denoted with •) were plotted against increasing guest concentration (mM). The calculated values for the Δδ are overlaid with the observed values giving K_a = 51 ± 3 M^{-1} (5.46%).

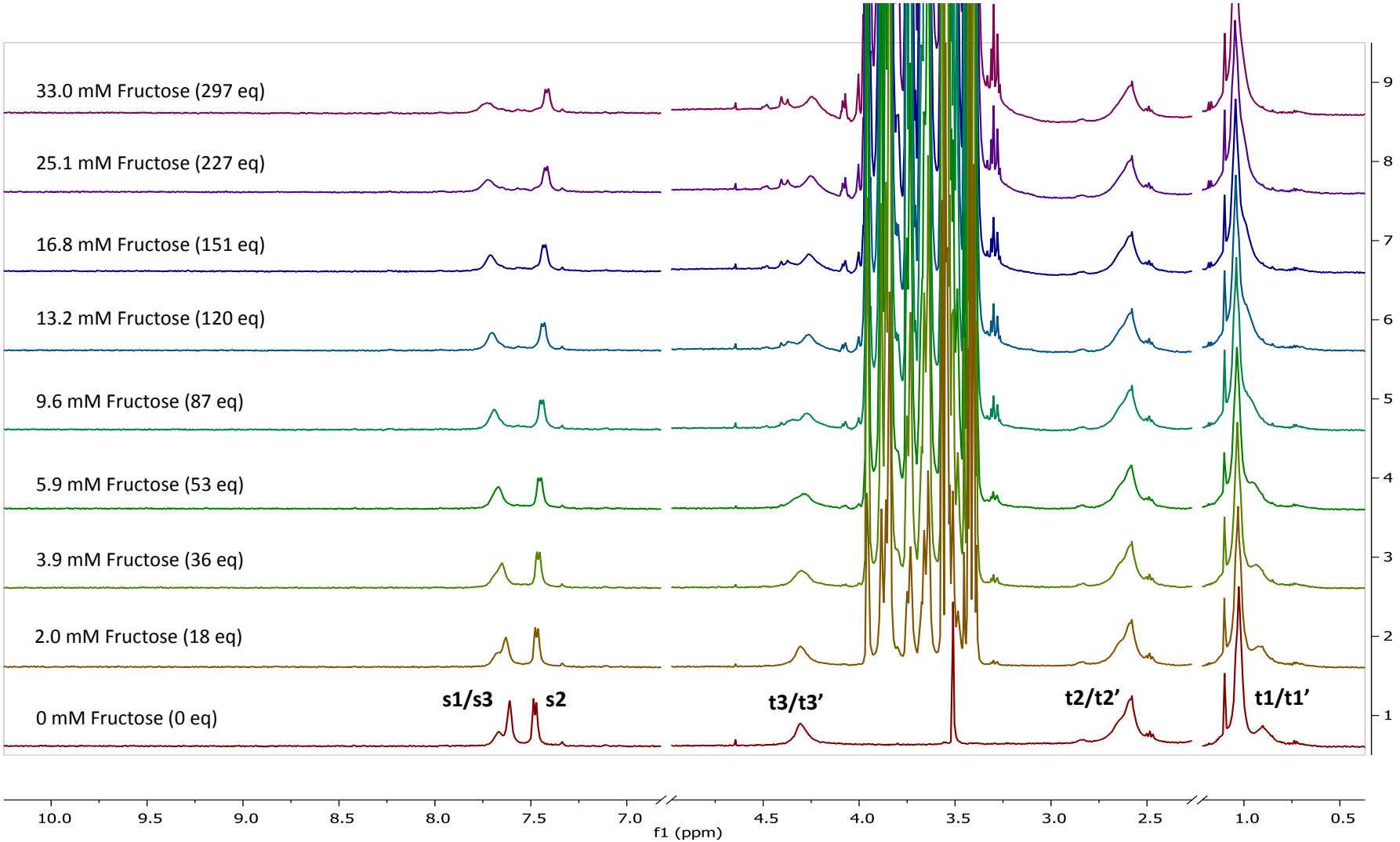


Figure 166 ^{1}H NMR spectra for receptor **3** (0.11 mM) titrated with a combined solution of D-fructose **20** (250 mM) and receptor **3** (0.11 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K

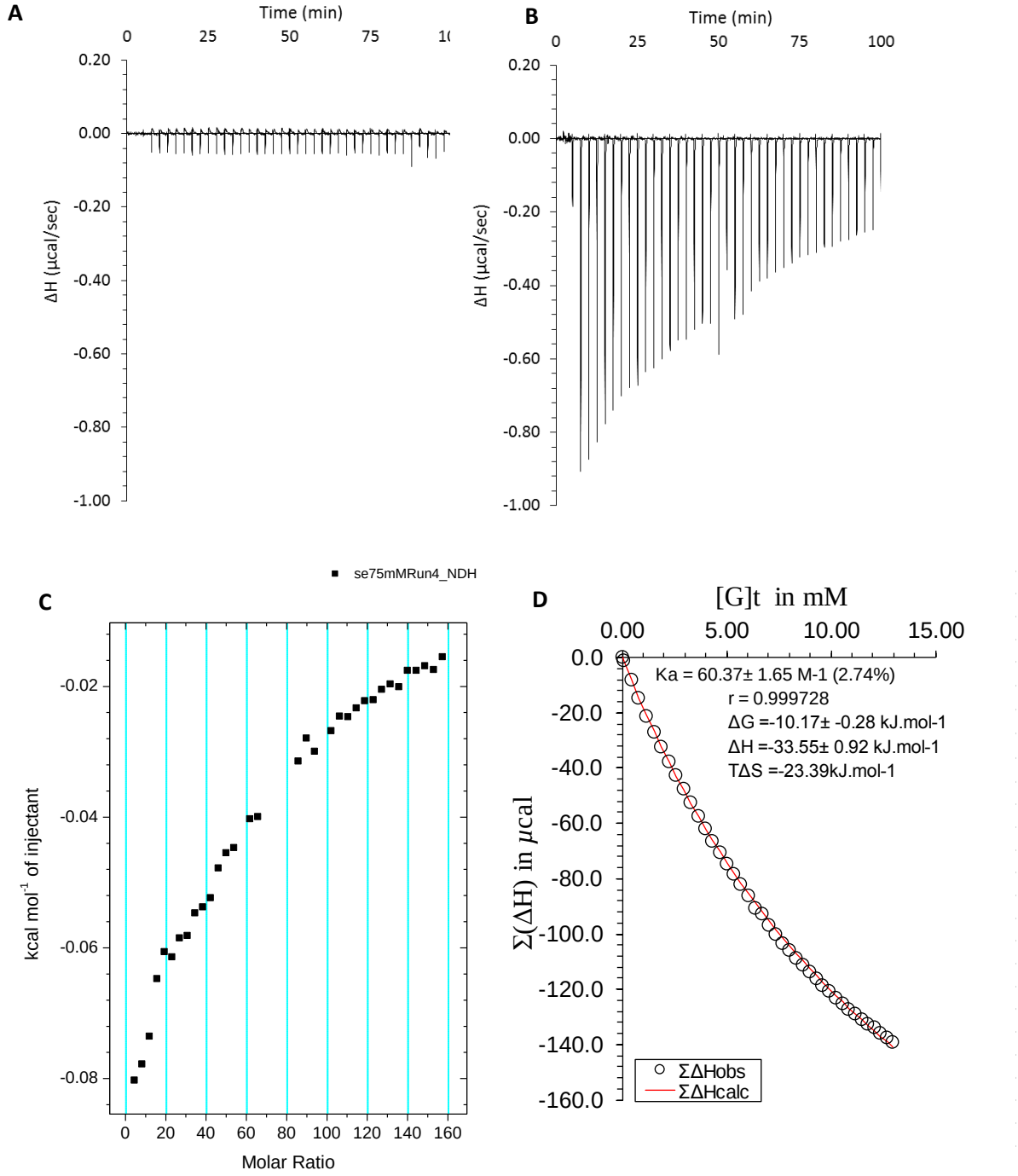


Figure 167 ITC binding results for receptor **3** (0.1 mM) titrated with D-fructose **20** (75 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 60.3 ± 1.6 M^{-1}).

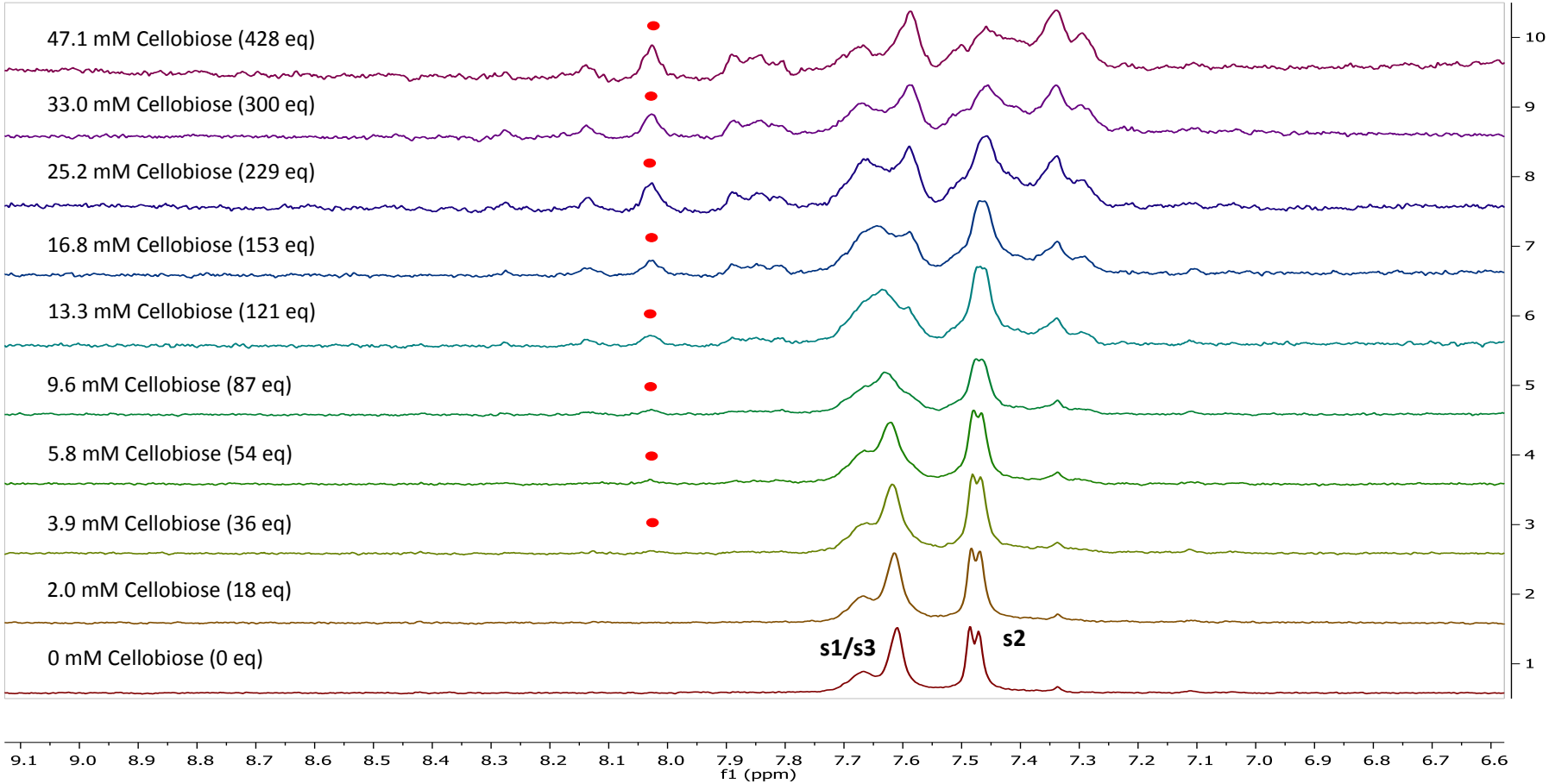


Figure 168 Partial ^{1}H NMR spectra for receptor **3** (0.11 mM) titrated with a combined solution of D-cellobiose **31** (250 mM) and receptor **3** (0.11 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with slow exchange on NMR timescale. Integrations of peak at 8.02 ppm (denoted with •) versus region 8.36-7.36 ppm were used to calculate the K_a (M^{-1}) at each point of addition (see Table S7), an average of these calculated values gives K_a = 31 ± 2.66 (9%).

Table 11 Calculated values for K_a when titrating cellobiose **31** (250 mM) against receptor **3** (0.11 mM) in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K.

	Calculated K_a / M^{-1}
	29.0
	31.2
	29.4
	34.9
Average K_a / M^{-1}	31.1
Std Dev (Error)	2.66 (9%)

The integrals of the peak at 8.02 ppm (denoted with •, see Figure 170) were measured relative to the region 8.36-7.36. These values were then compared to that at the end of the titration, when all receptor is assumed to be saturated with guest. Based on this assumption, K_a could be calculated from each point in the titration. An average K_a = 31.1 M^{-1} ± 2.66 (9%) was obtained in agreement with results obtained from ITC studies. Not all the spectra were used due to baseline distortions resulting from the large excess of guest present.

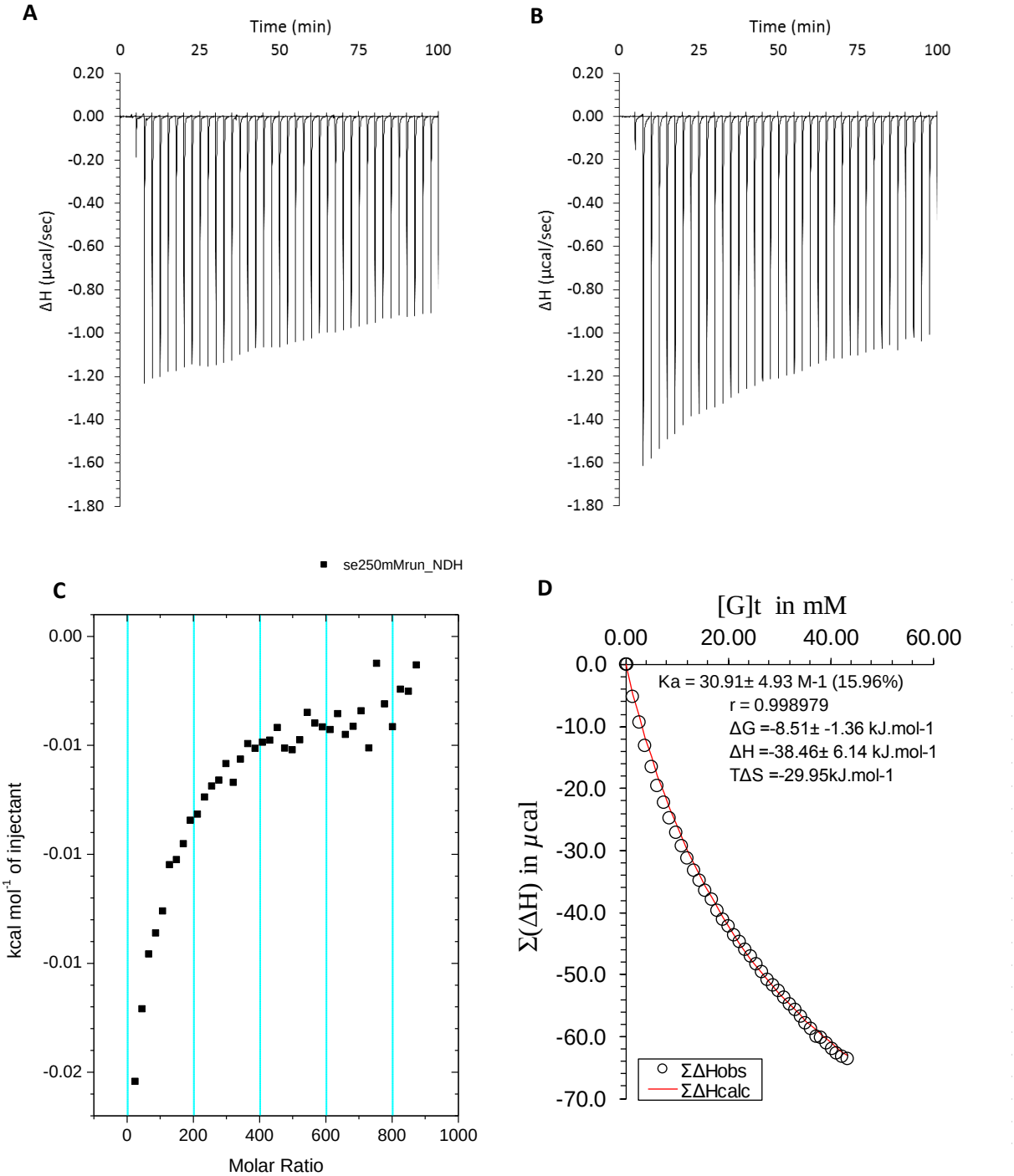


Figure 169 ITC binding results for receptor **3** (0.06 mM) titrated with D-cellobiose **31** (250 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (K_a = 30.9 ± 4.9 M^{-1}).

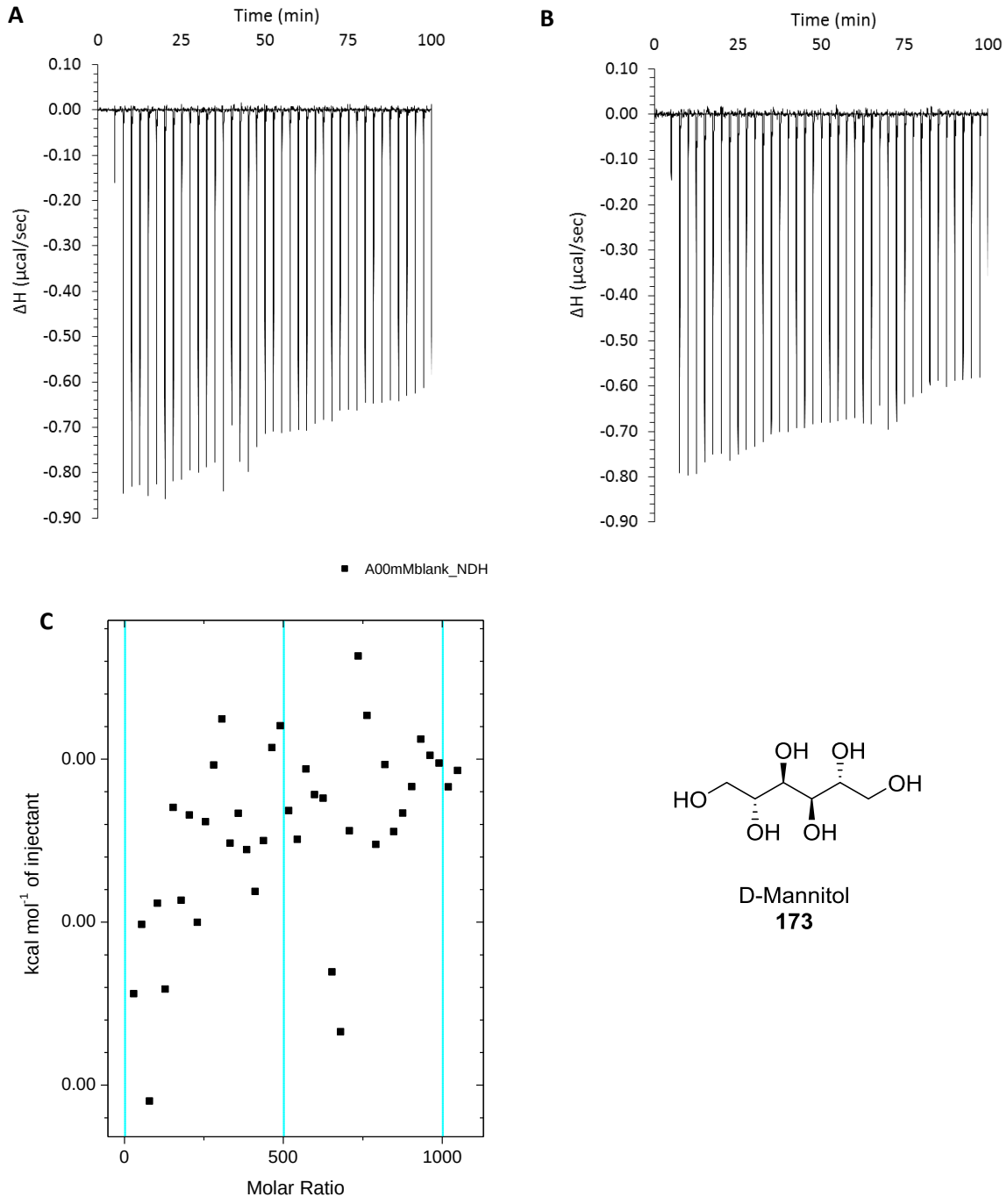


Figure 170 ITC binding results for receptor **3** (0.1 mM) titrated with D-mannitol **173** (500 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio

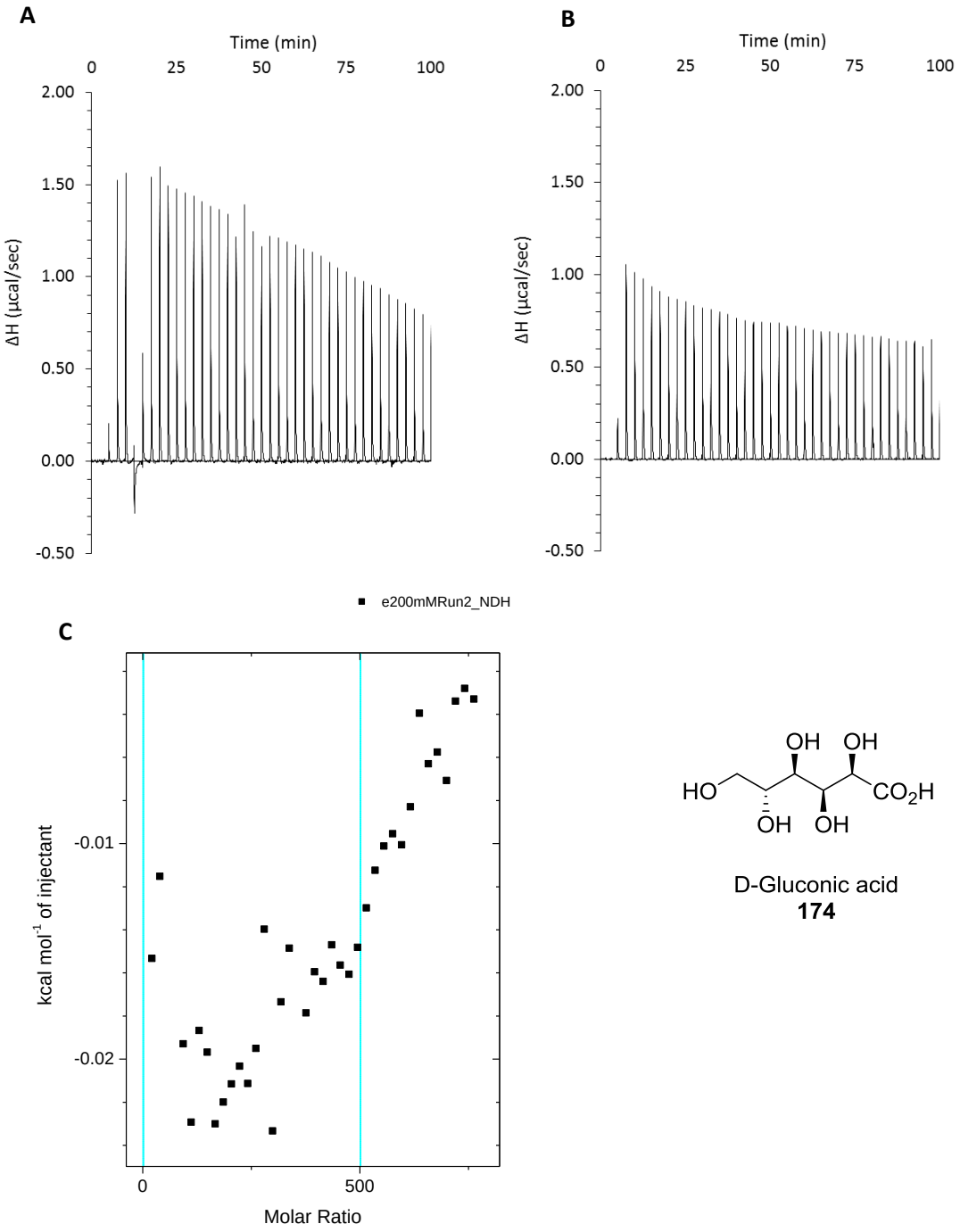


Figure 171 ITC binding results for receptor **3** (0.06 mM) titrated with D-gluconic acid **174** (200 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of substrate into water); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio.

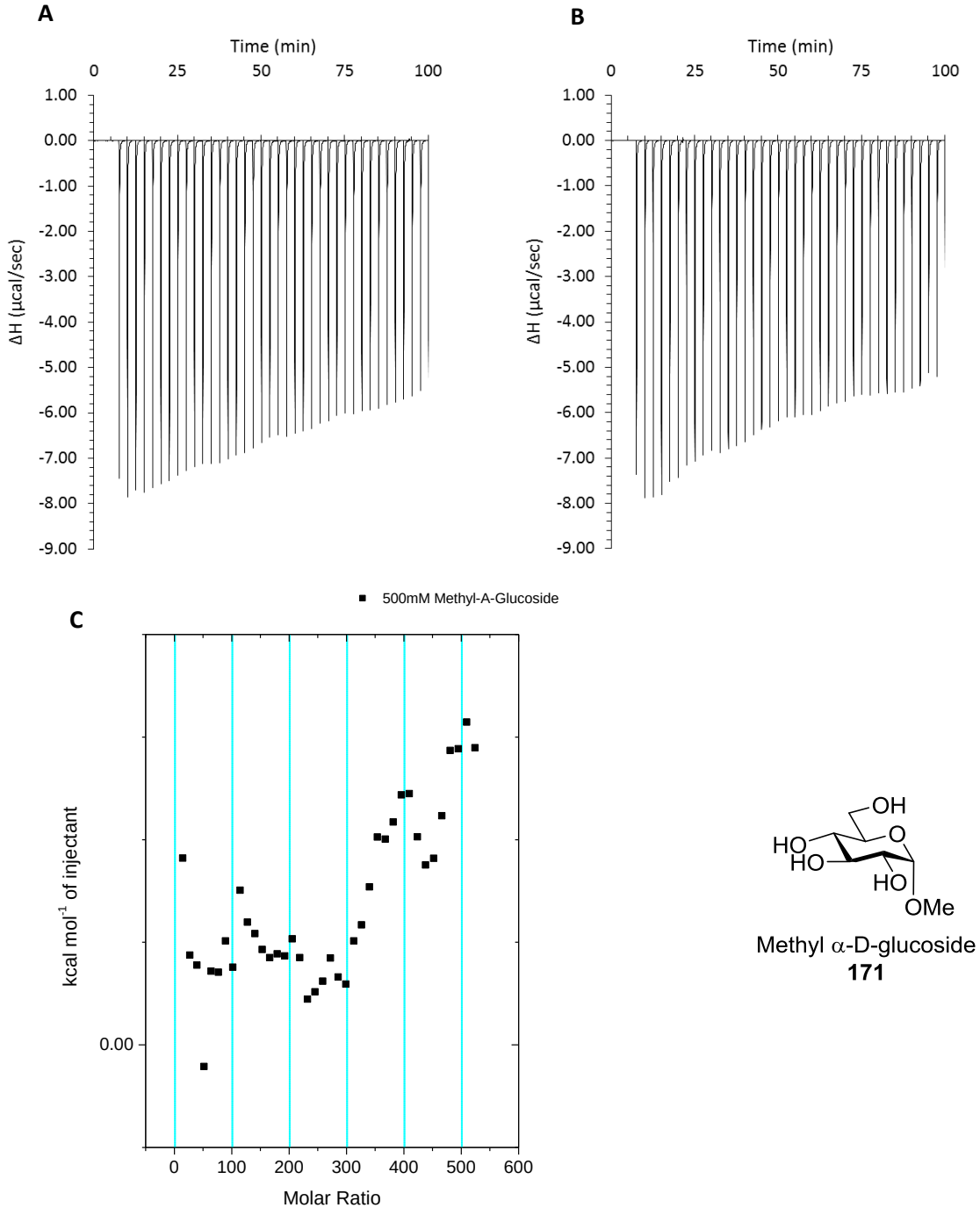


Figure 172 ITC binding results for receptor **3** (0.06 mM) titrated with methyl α-D-glucoside **171** (500 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**) and C) shows the plotted change in enthalpy vs molar ratio.

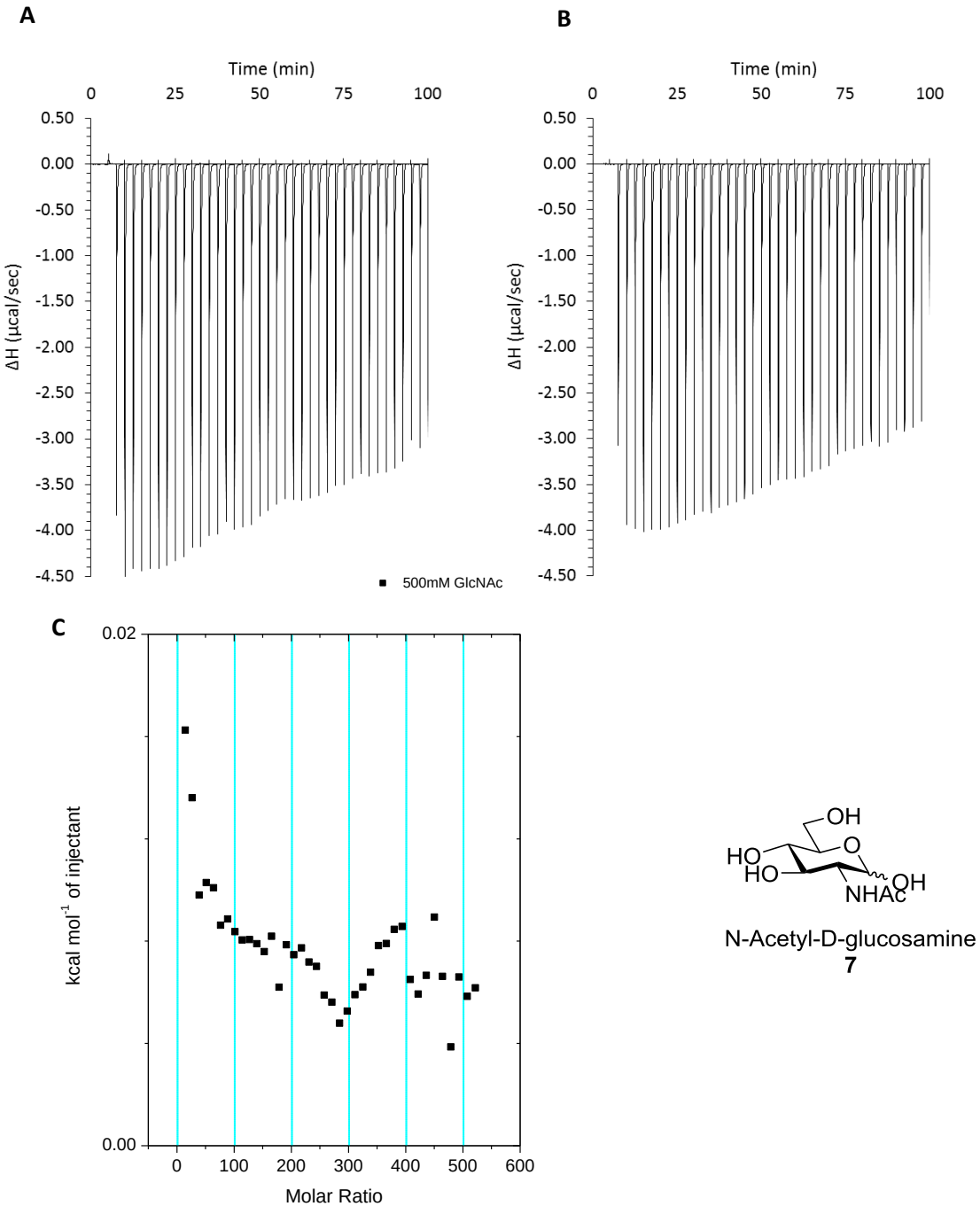


Figure 173 ITC binding results for receptor **3** (0.06 mM) titrated with *N*-acetyl-D-glucosamine **7** (498 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**) and C) shows the plotted change in enthalpy vs molar ratio.

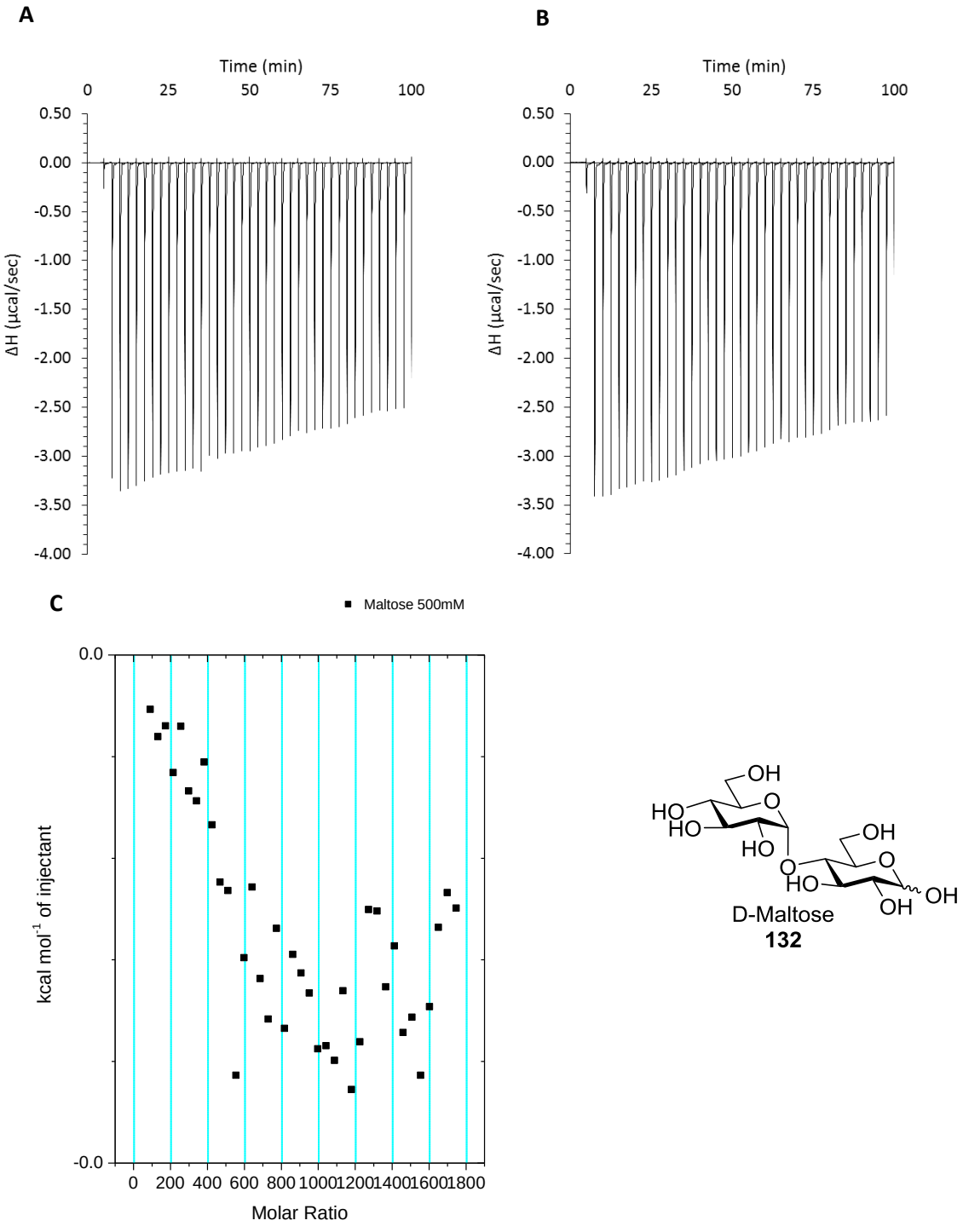


Figure 174 ITC binding results for receptor **3** (0.06 mM) titrated with maltose **132** (500 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of sugar into water); B) shows the titration (sugar into receptor **3**) and C) shows the plotted change in enthalpy vs molar ratio.

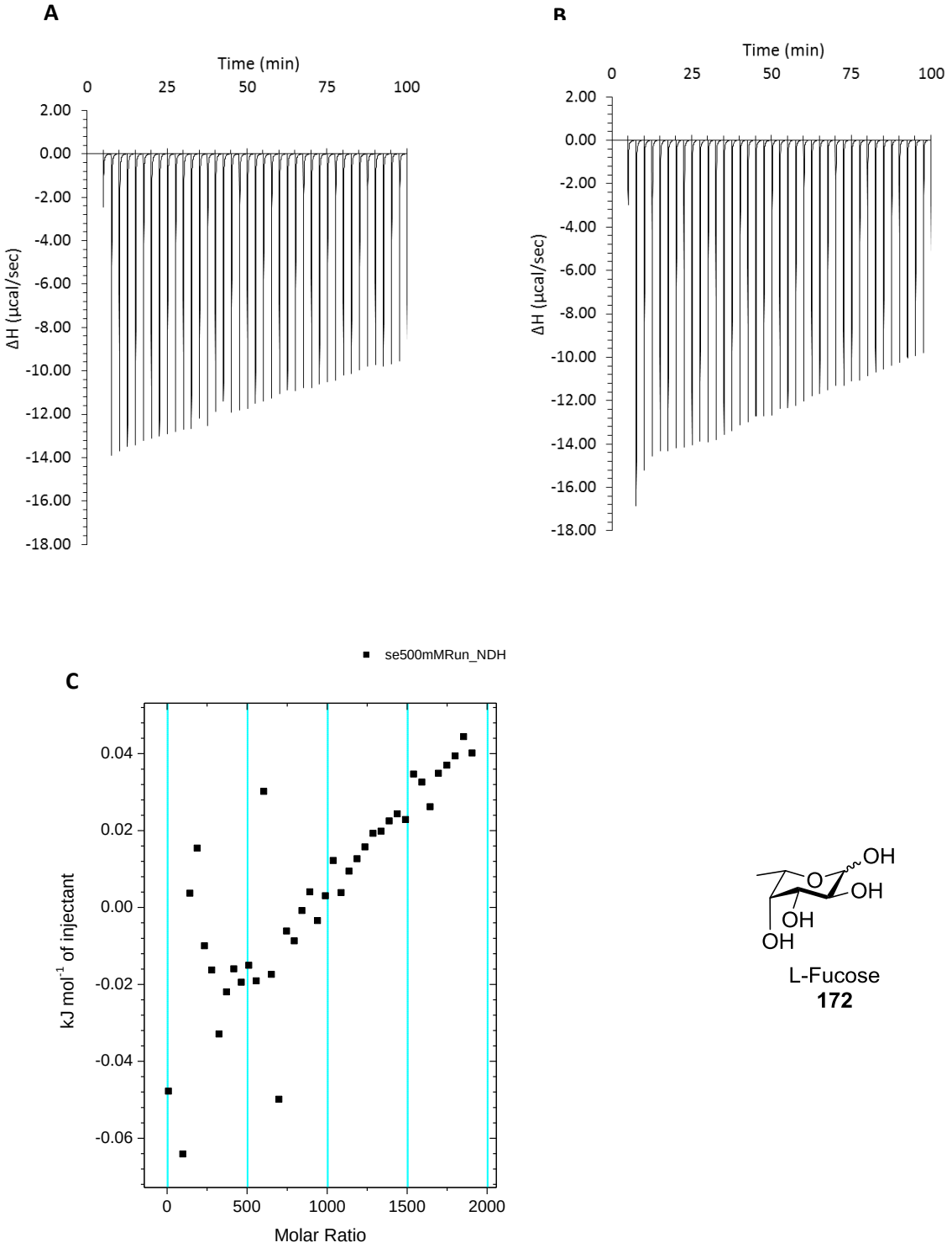


Figure 175 ITC binding results for receptor **3** (0.06 mM) titrated with L-fucose **172** (500 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of substrate into water); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio.

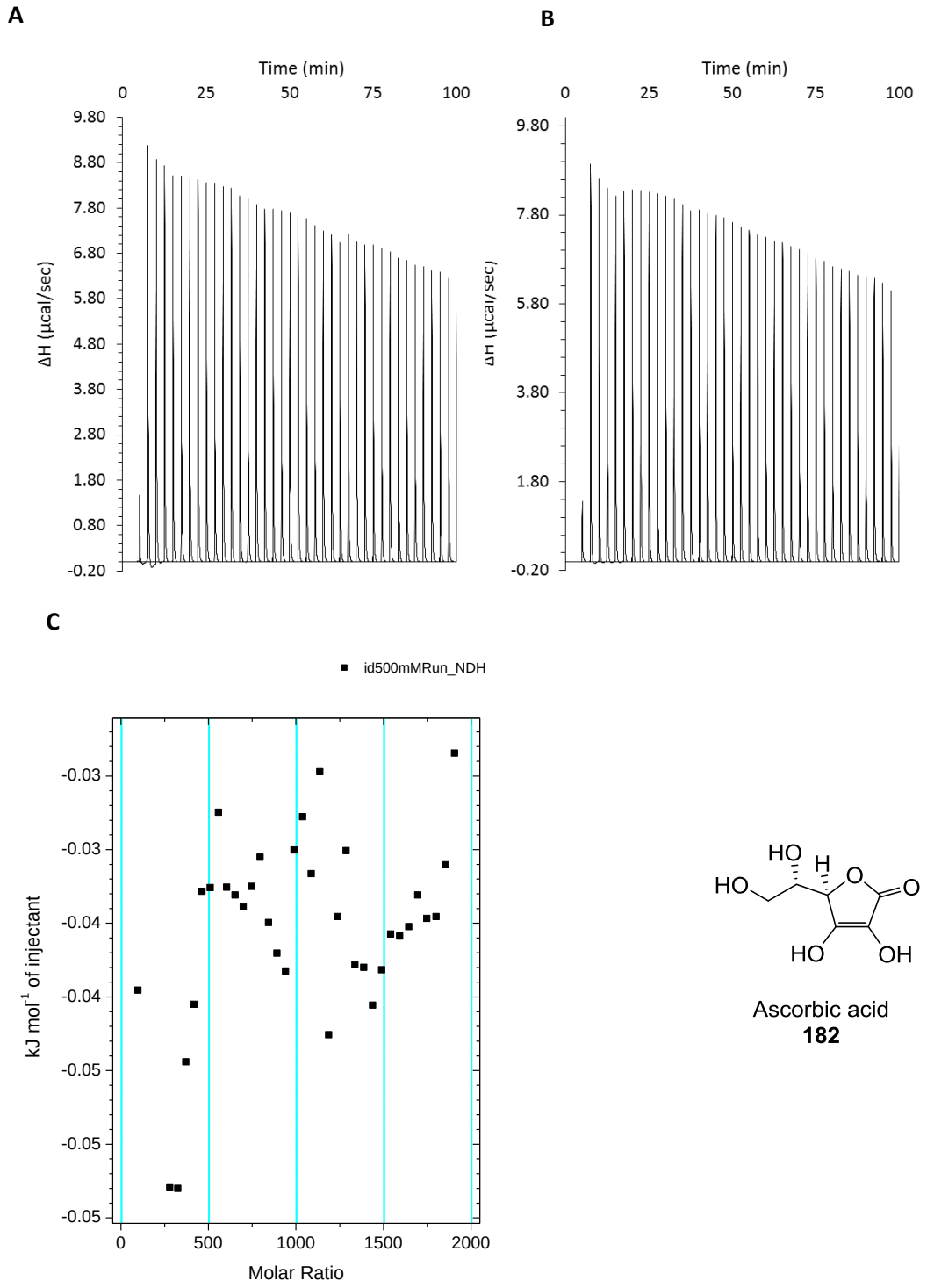


Figure 176 ITC binding results for receptor **3** (0.06 mM) titrated with ascorbic acid **182** (500 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of substrate into water); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio.

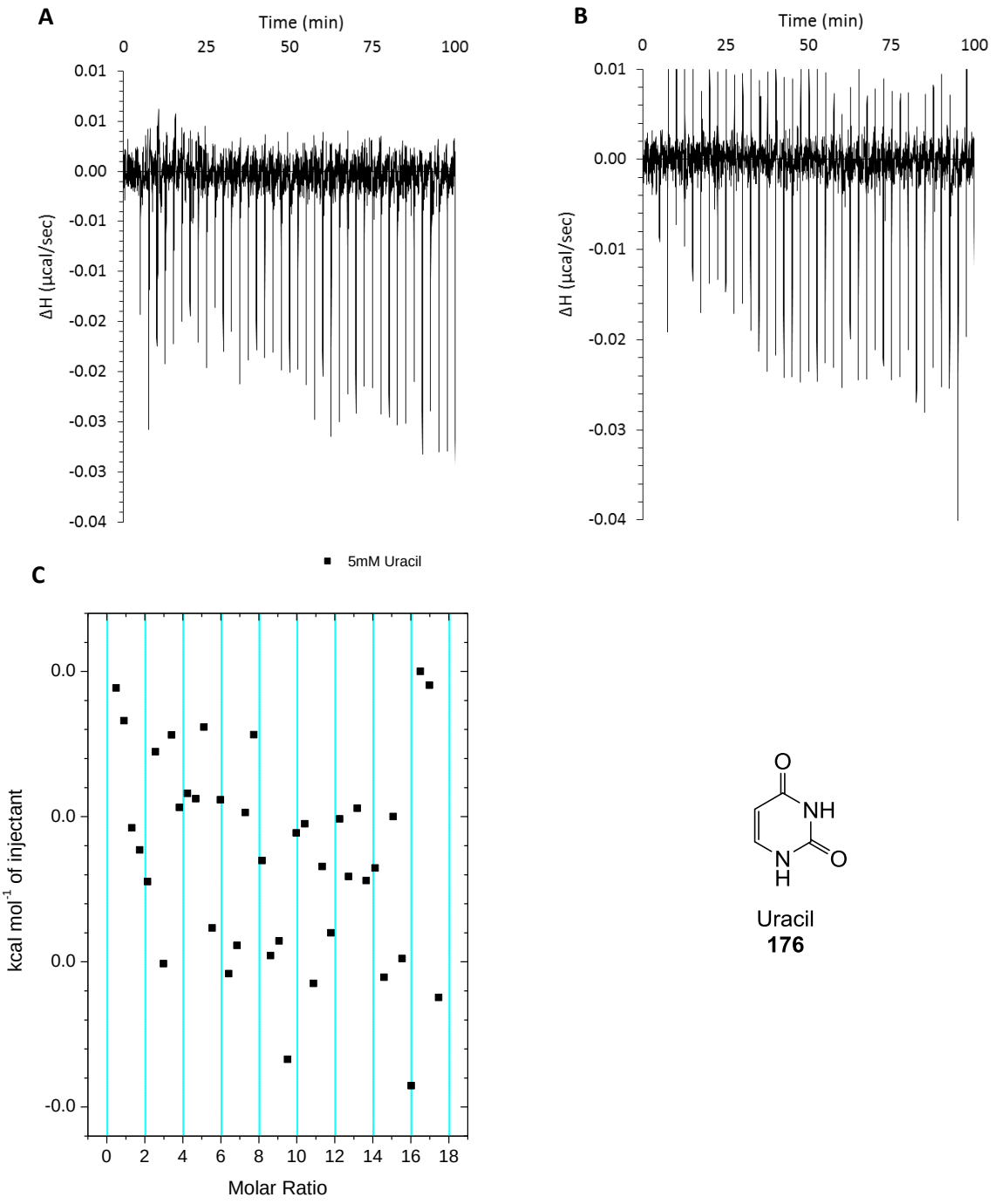


Figure 177 ITC binding results for receptor **3** (0.06 mM) titrated with uracil **176** (5 mM, limit of solubility) in 10 mM phosphate buffered saline (pH 7.4), in which: A) shows the blank run (addition of uracil into water); B) shows the titration (uracil into receptor **3**) and C) shows the plotted change in enthalpy vs molar ratio.

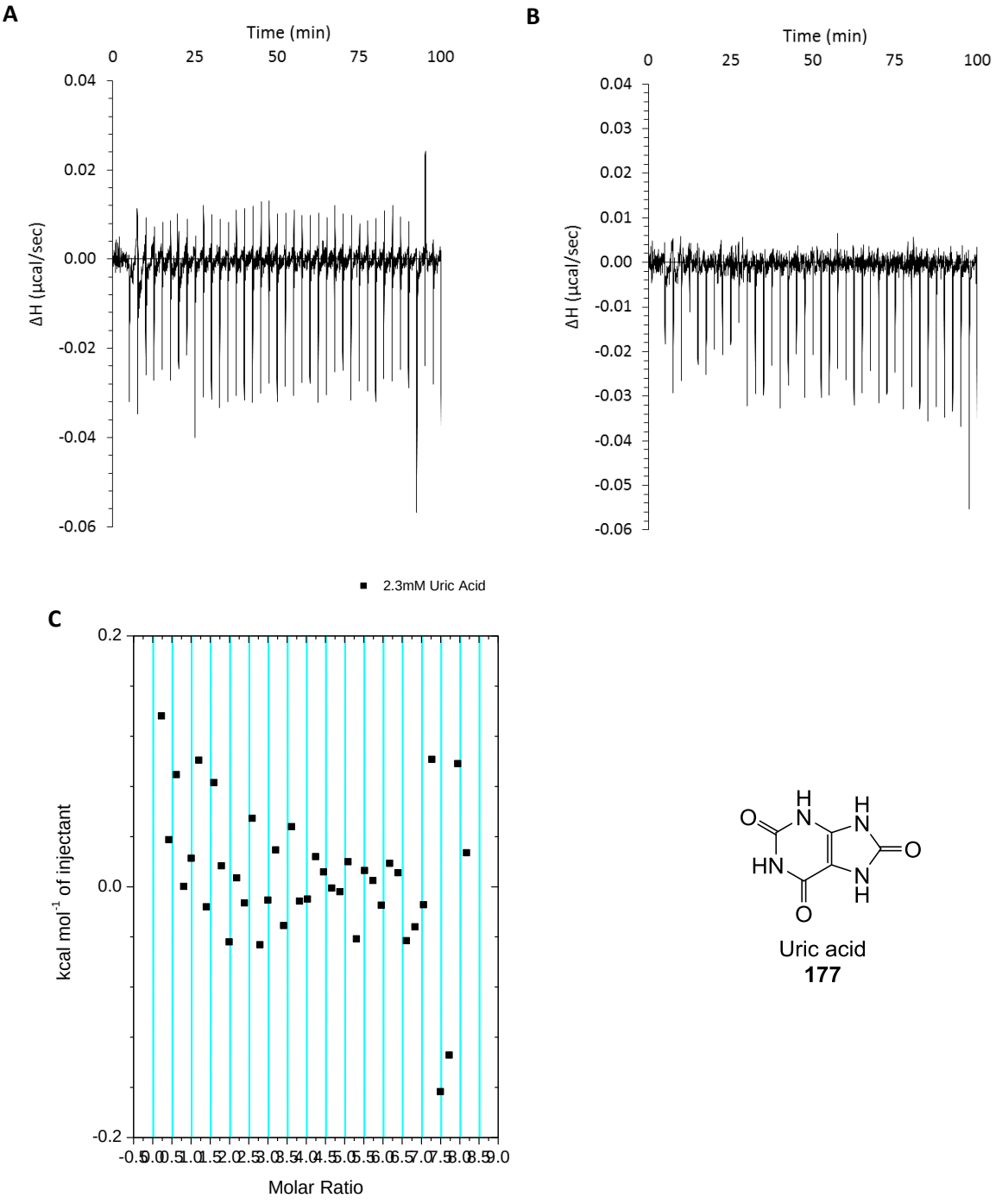


Figure 178 ITC binding results for receptor **3** (0.06 mM) titrated uric acid **177** (2.34 mM, limit of solubility) in 10 mM phosphate buffered saline (pH 7.4), in which: A) shows the blank run (addition of uric acid into water); B) shows the titration (uric acid into receptor **3**) and C) shows the plotted change in enthalpy vs molar ratio.

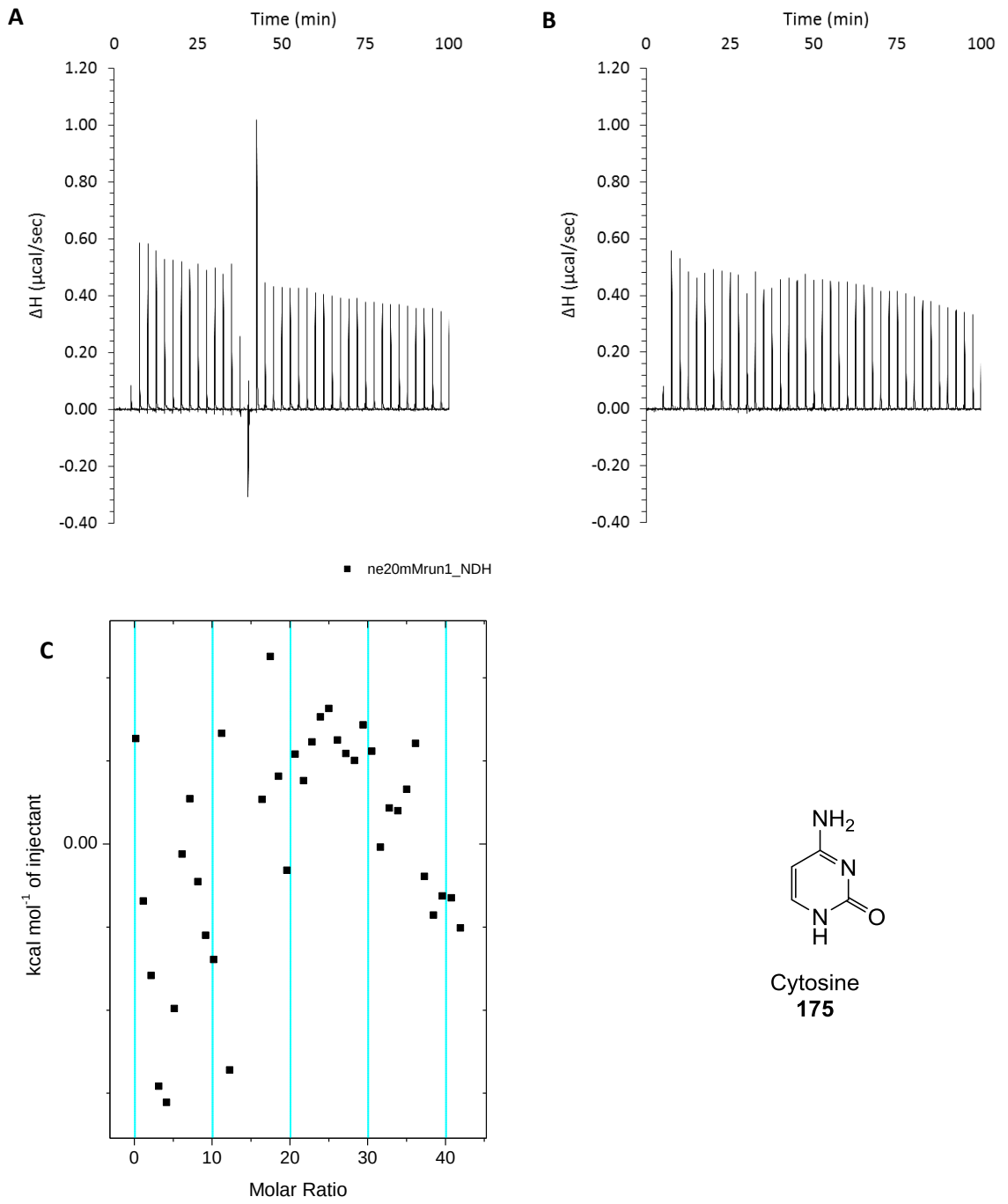


Figure 179 ITC binding results for receptor **3** (0.1 mM) titrated with cytosine **175** (20 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio.

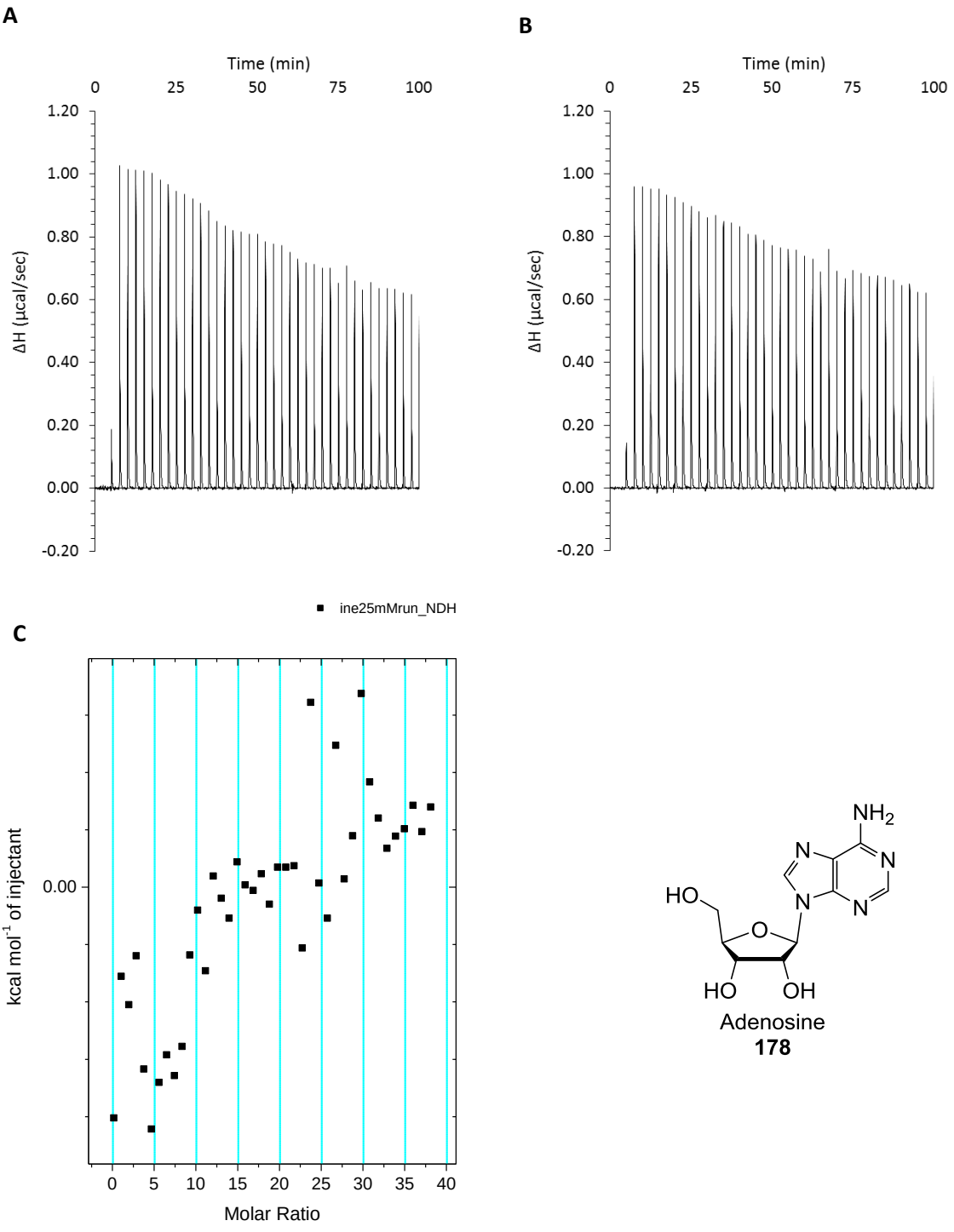


Figure 180 ITC binding results for receptor **3** (0.1 mM) titrated with adenosine **178** (500 mM) in 10 mM phosphate buffer solution (pH 7.4) in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio.

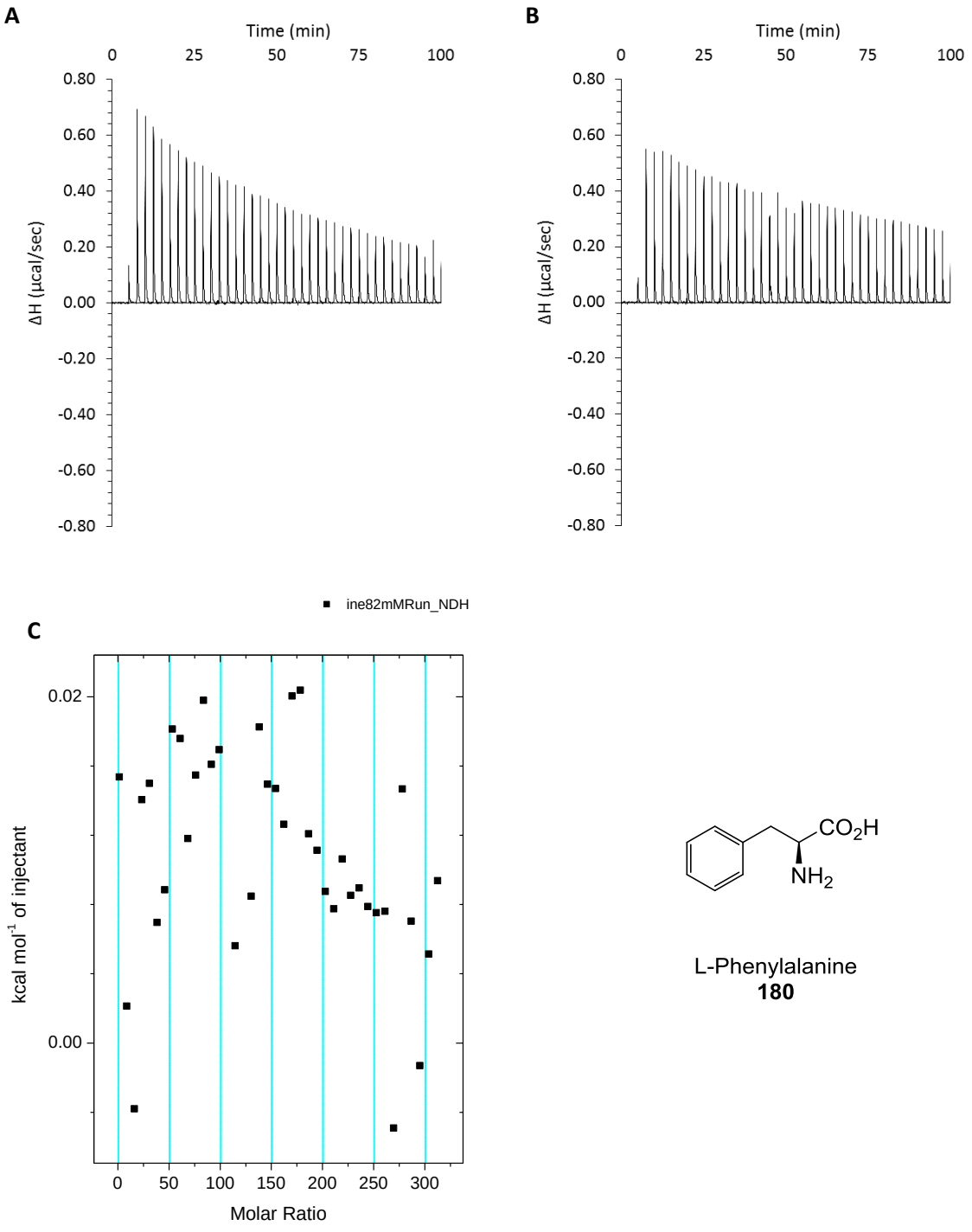


Figure 181 ITC binding results for receptor **3** (0.06 mM) titrated with L-phenylalanine **180** (82 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio.

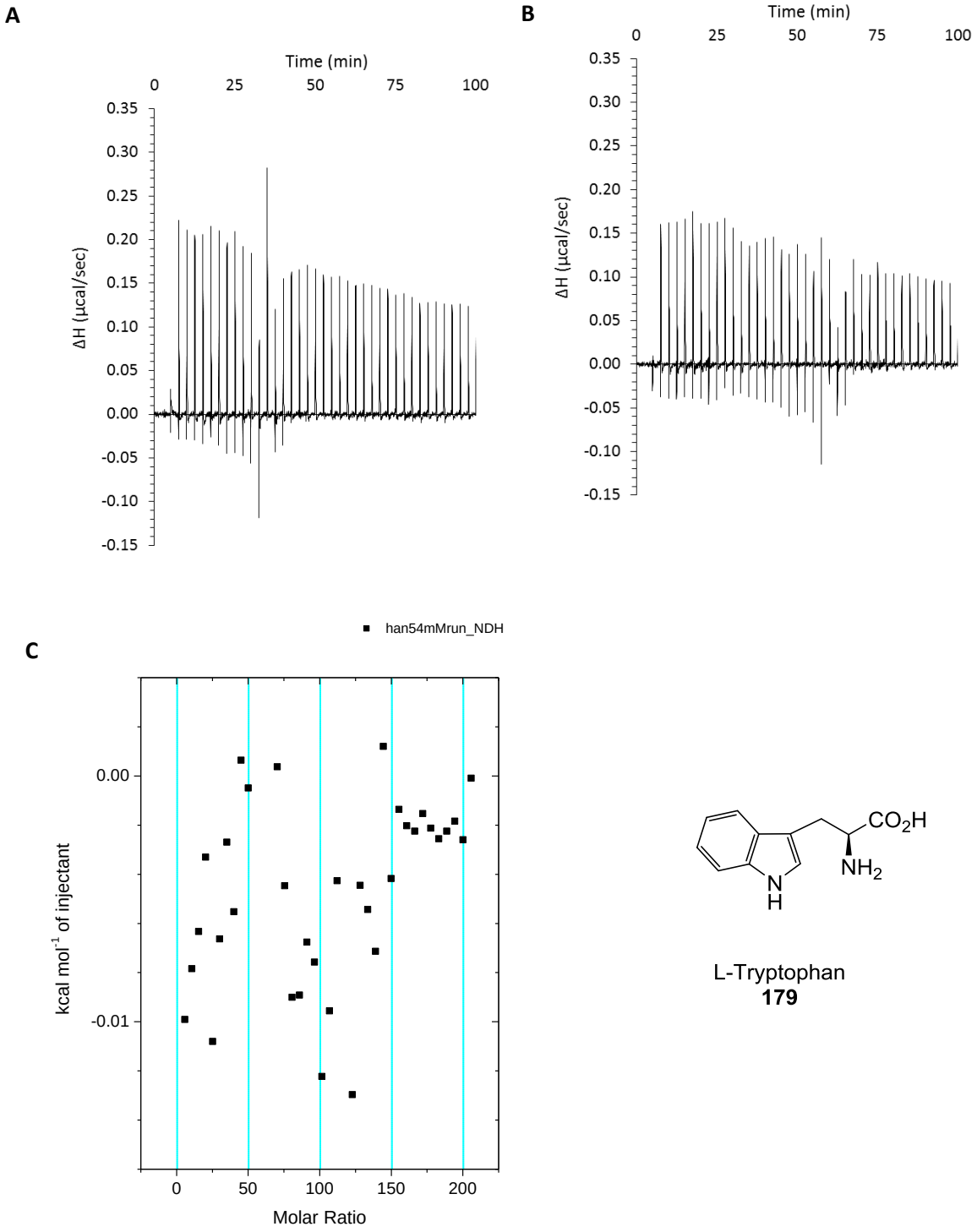


Figure 182 ITC binding results for receptor **3** (0.06 mM) titrated with L-tryptophan **179** (54 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of substrate into medium); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio.

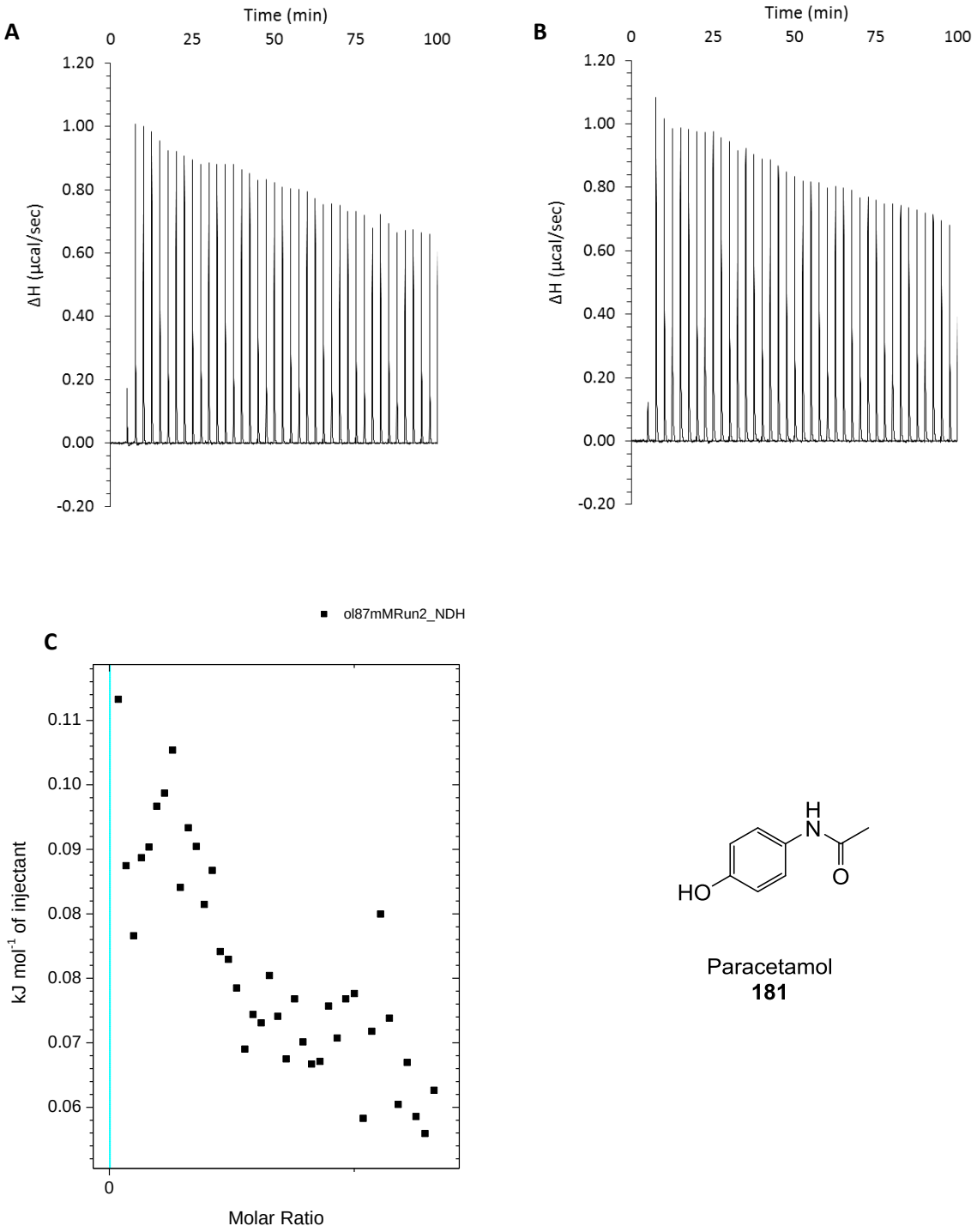


Figure 183 ITC binding results for receptor **3** (0.06 mM) titrated with paracetamol **181** (87 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of substrate into water); B) shows the titration (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio.

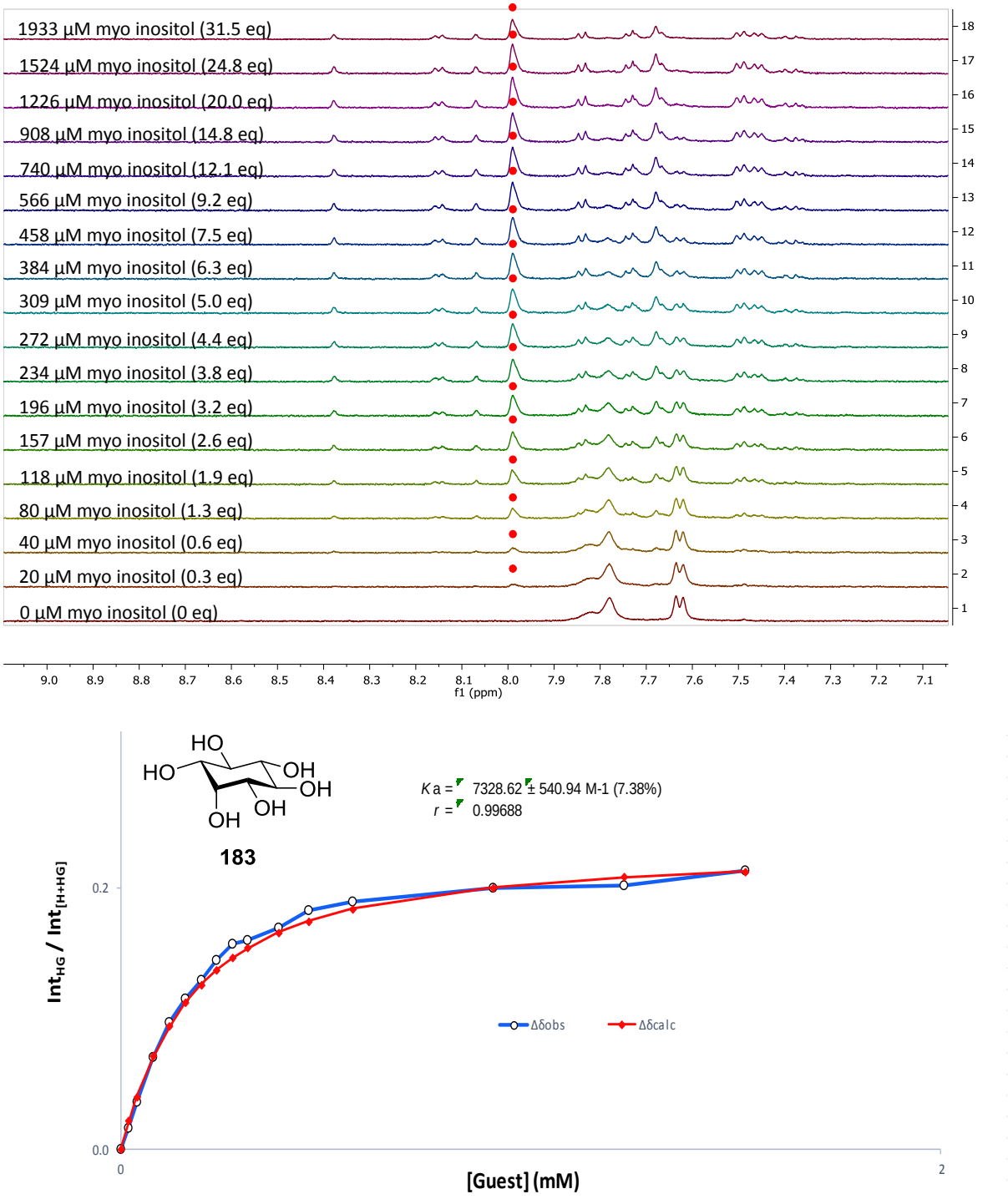


Figure 184 ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.07 mM) titrated with a combined solution of *myo*-inositol **183** (10 mM) and receptor **3** (0.07 mM), in D_2O buffered with 10 mM phosphate buffer solution (pH 7.4) at 298 K. Spectra imply binding with slow exchange on NMR timescale. Integrations of peak at 7.98 ppm (denoted with •) versus region 8.39-7.32 ppm were plotted against guest concentration (mM). The calculated values for the integrals are overlaid with the observed values, giving K_a = 7328 ± 540 M^{-1} (7.4%)

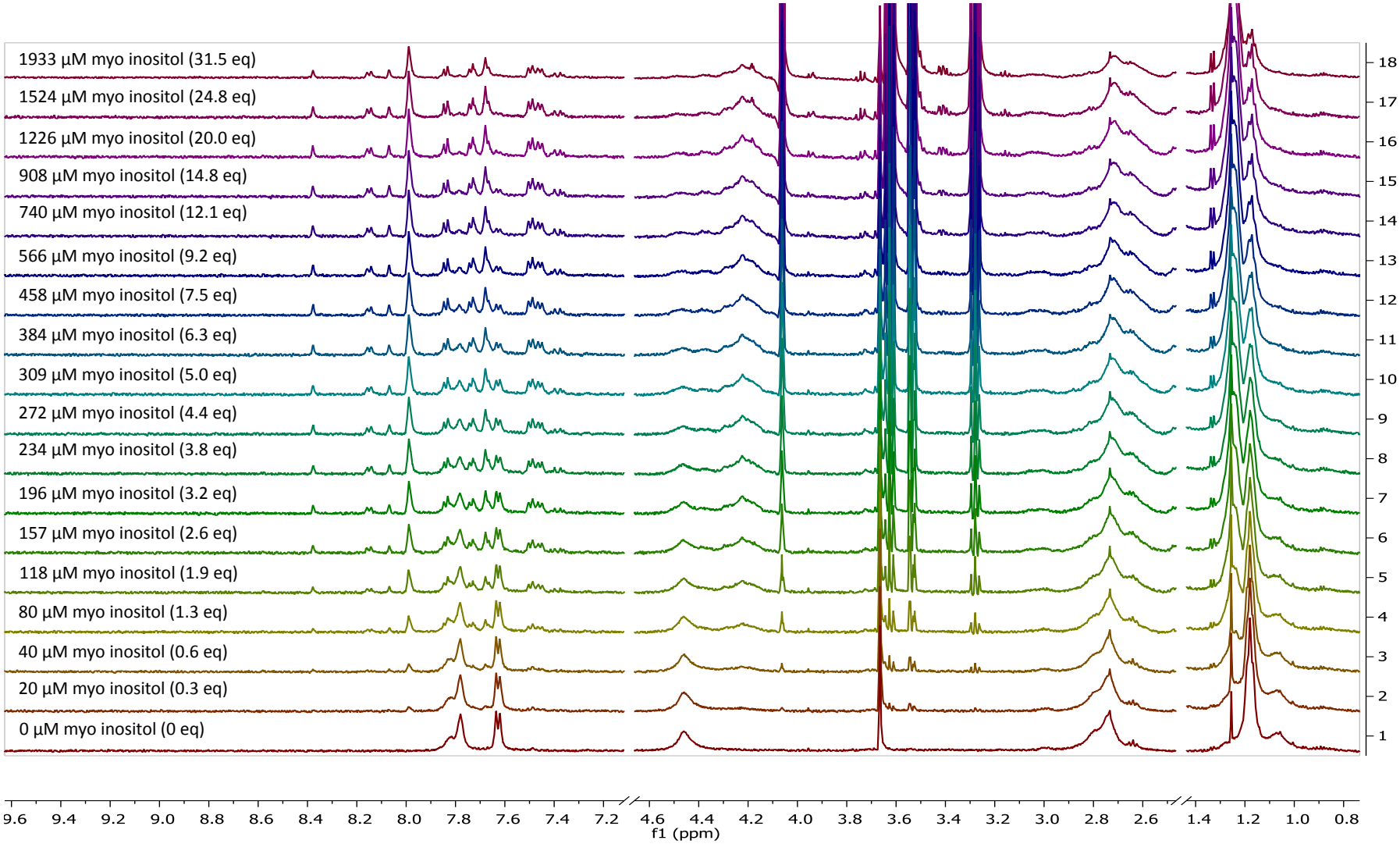


Figure 185 ^{1}H NMR spectra for receptor **3** (0.07 mM) titrated with a combined solution of *myo*-inositol **183** (10 mM) and receptor **3** (0.07 mM), in D_2O buffered with 10 mM phosphate buffer solution (pH 7.4) at 298 K.

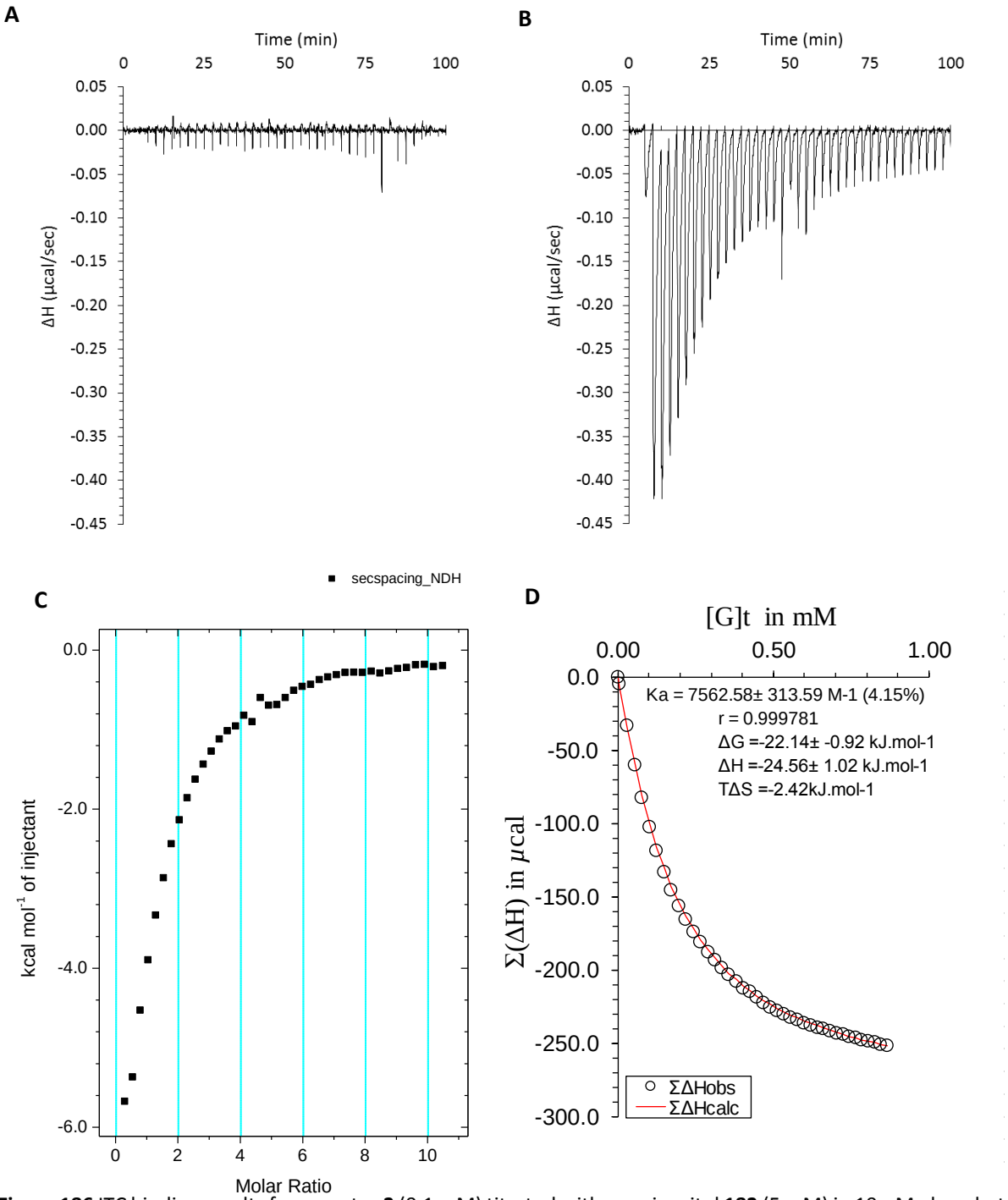


Figure 186 ITC binding results for receptor **3** (0.1 mM) titrated with *myo*-inositol **183** (5 mM) in 10mM phosphate buffer solution (pH 7.4) in which: A) shows the blank ITC run (addition of substrate into medium); B) shows the actual run (substrate into receptor **3**); C) shows the plotted change in enthalpy vs molar ratio; and D) shows the fit calculated using an Excel spreadsheet (Ka = 7563 ± 313 M^{-1}).

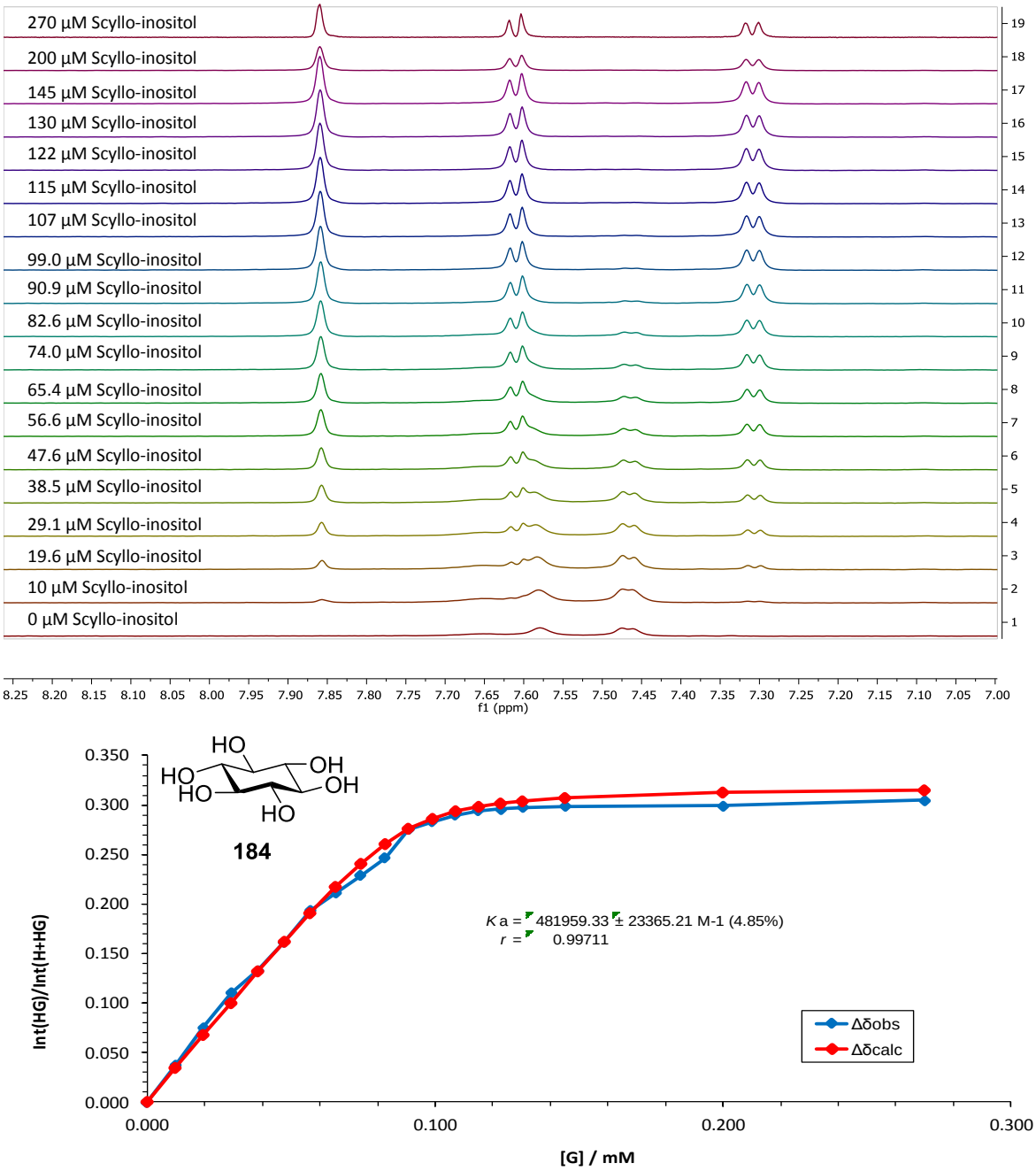


Figure 187 Partial ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **3** (0.09 mM) titrated with a combined solution of *scyllo*-inositol **184** (1 mM) and receptor **3** (0.09 mM), in D_2O with 10 mM phosphate buffer (pH 7.4) at 298 K. Spectra imply binding with slow exchange on NMR timescale. Integrations of the peak at δ 7.85 ppm versus the region δ 7.9-7.25 ppm were plotted against *scyllo*-inositol **184** concentration (mM). The calculated values for the integrals are overlaid with the observed values, giving K_a = 482,000 ± 23,400 M^{-1} (4.9%).

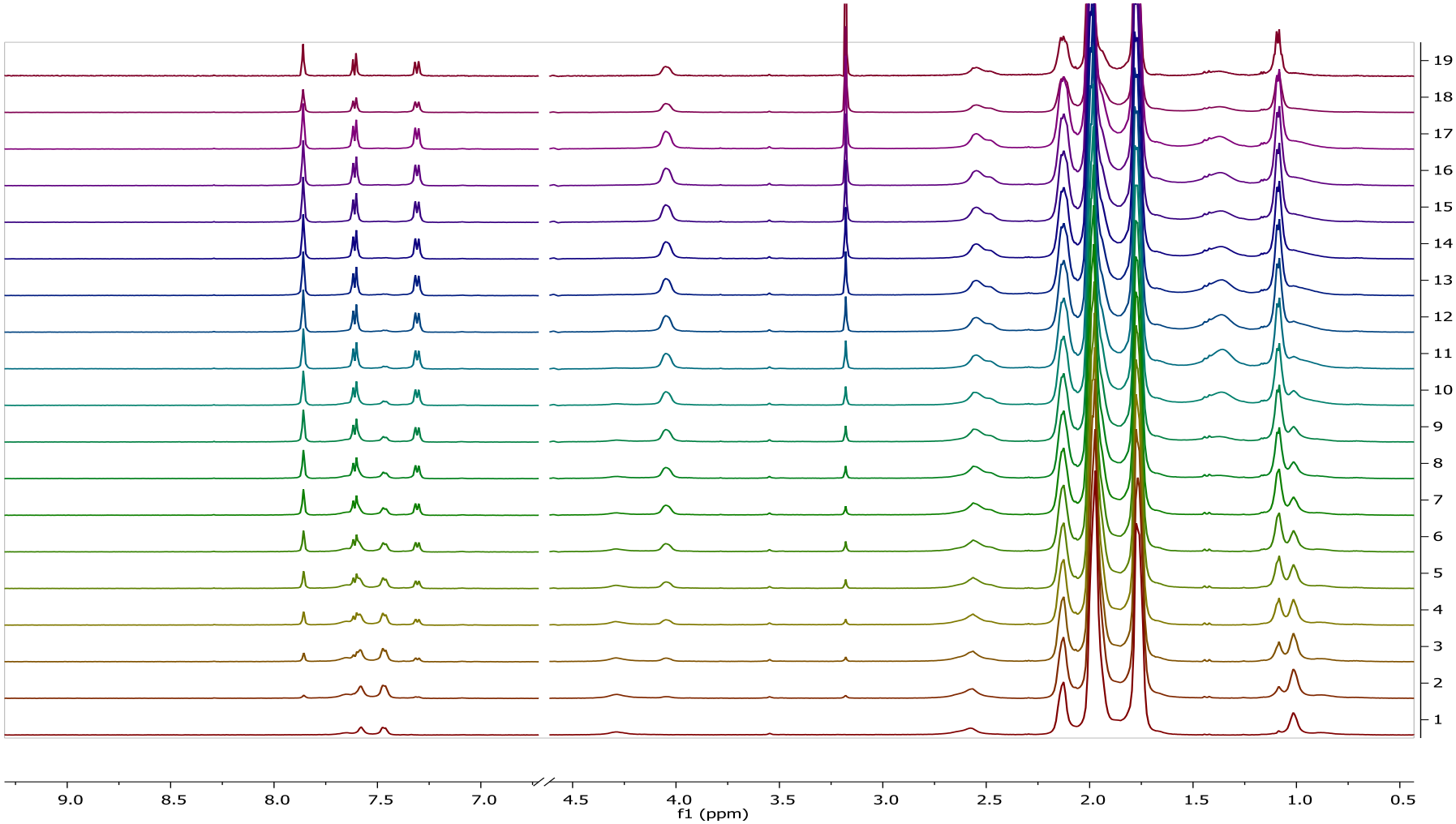


Figure 188 ^{1}H NMR spectra for receptor **3** (0.09 mM) titrated with a combined solution of *scyllo*-inositol **184** (1 mM) and receptor **3** (0.09 mM), in D_2O buffered with 10 mM phosphate buffer solution (pH 7.4) at 298 K.

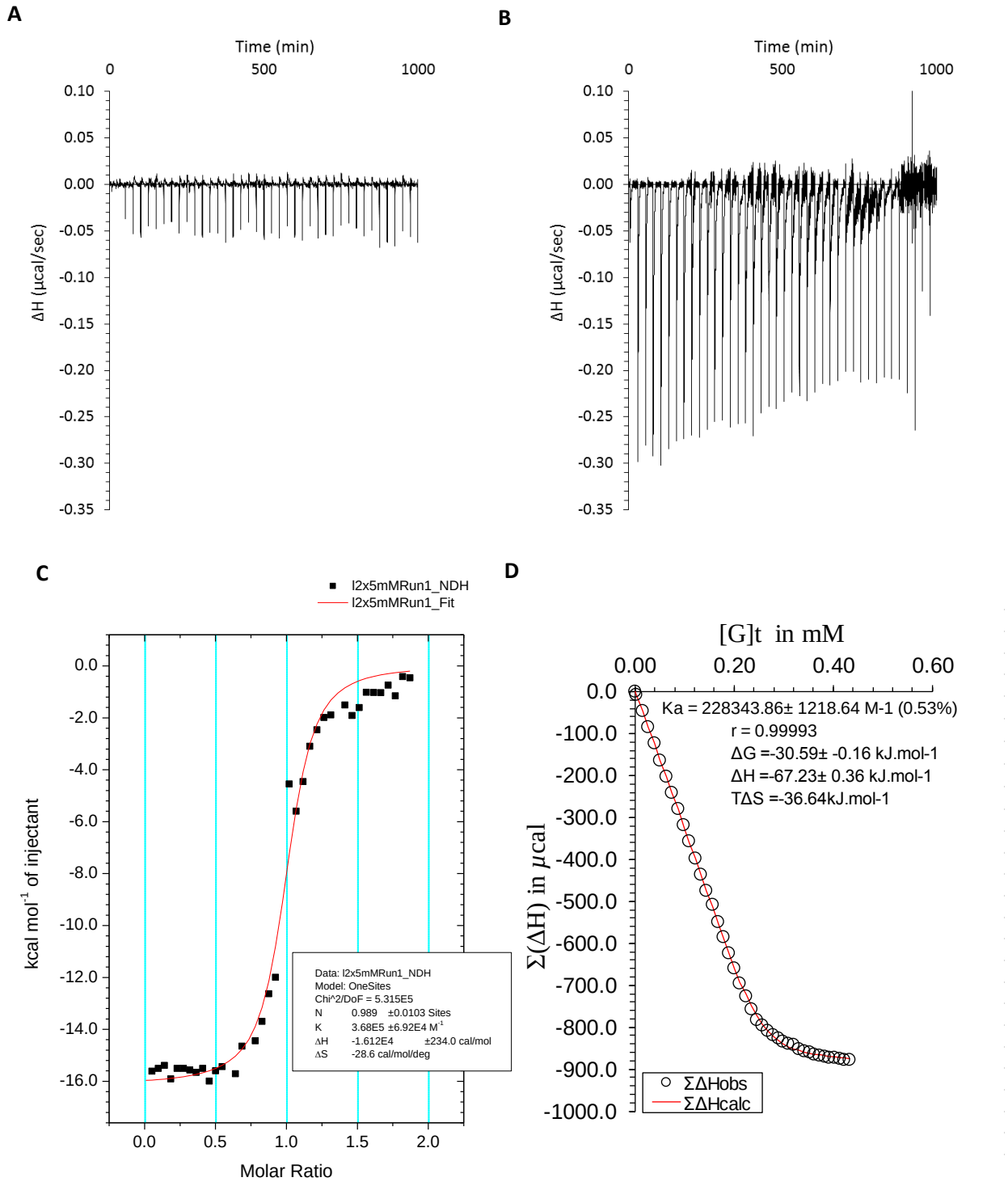
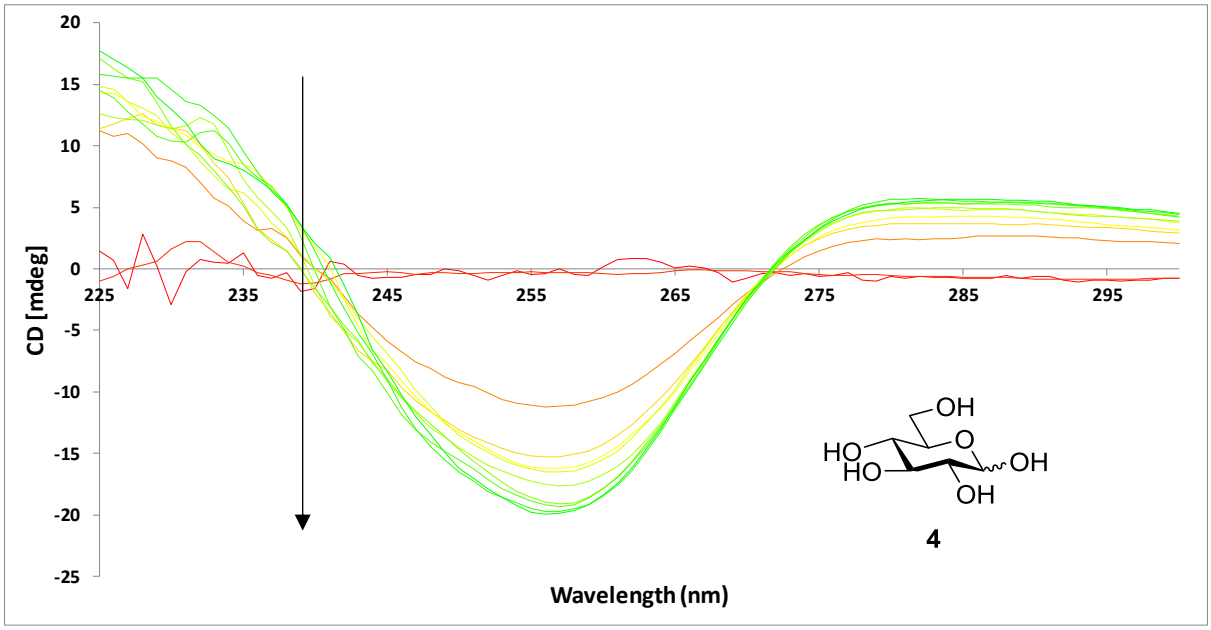


Figure 189 ITC binding results for receptor **3** (0.2 mM) titrated with *scyllo*-inositol **184** (2.5 mM) in 10 mM phosphate buffer solution (pH 7.4), in which: A) shows the blank run (addition of substrate into water); B) shows the titration (substrate into receptor); C) shows the plotted change in enthalpy vs molar ratio and the fit using the MicroCal software (K_a = 368,000± 69,200 M^{-1}); and D) shows the fit calculated using an Excel spreadsheet (K_a = 228,000 ± 1210 M^{-1}).

9.2.4 Binding studies for triethylbenzene receptor (3) using circular dichroism (CD)



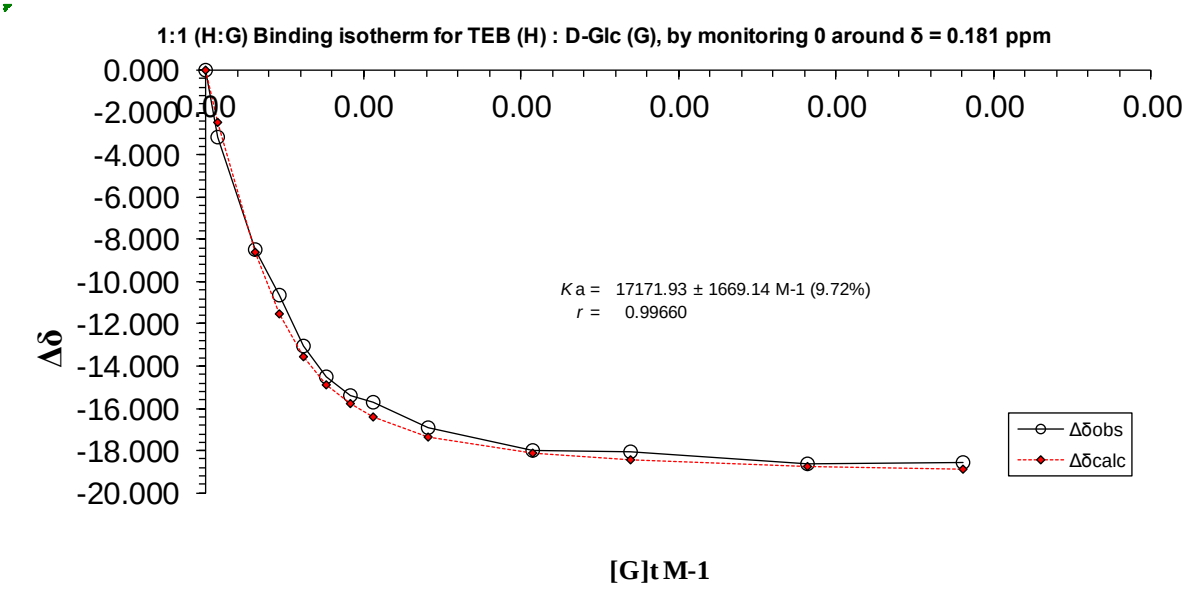
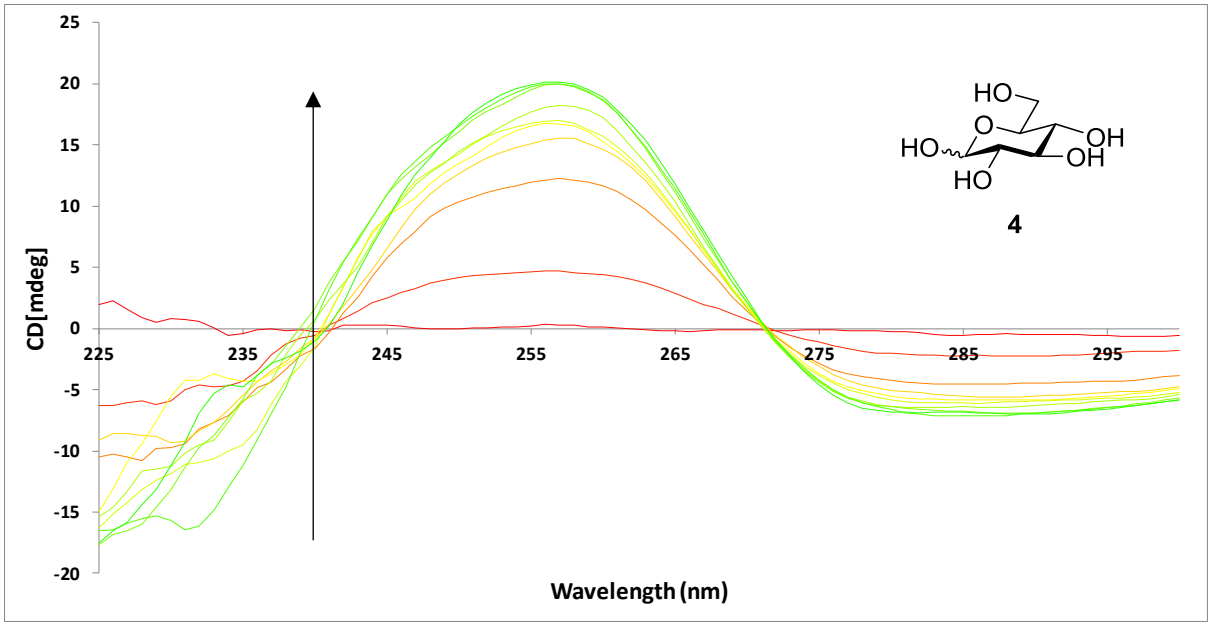


Figure 190 CD spectra (top) and binding analysis curve (bottom) for receptor **3** (0.25 mM) titrated with a combined solution of D-glucose **4** (10 mM) and receptor **3** (0.25 mM), in water with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing guest concentration (mM). The calculated values for the ΔCD are overlaid with the observed values giving K_a = 17,200 ± 1700 M^{-1} (9.7%).



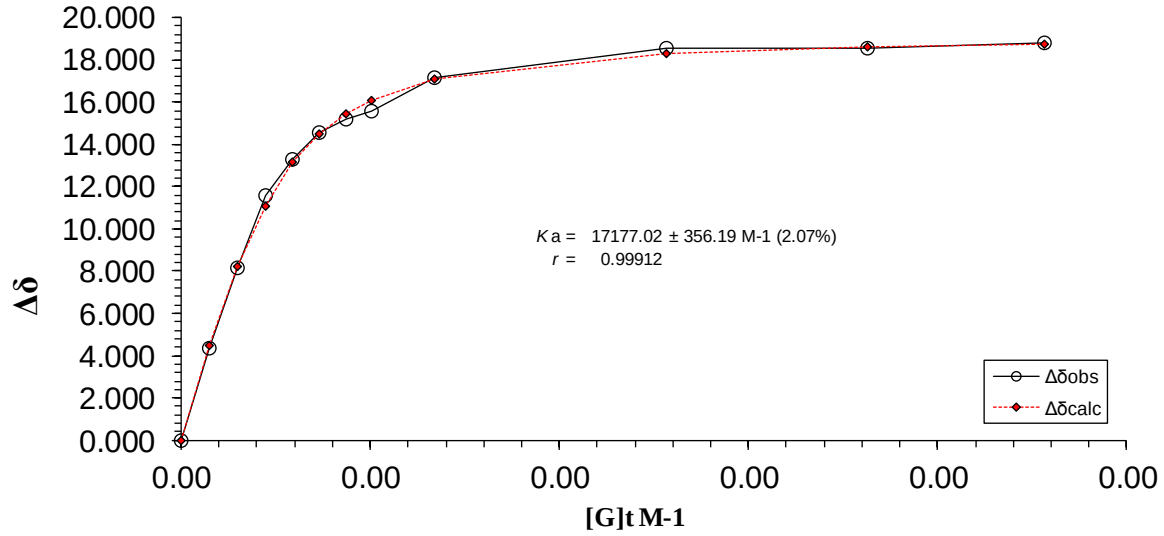


Figure 191 CD spectra (top) and binding analysis curve (bottom) for receptor **3** (0.25 mM) titrated with a combined solution of L-glucose **187** (10 mM) and receptor **3** (0.25 mM), in water with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing guest concentration (mM). The calculated values for the ΔCD are overlaid with the observed values giving K_a = 17,200 ± 360 M^{-1} (2.1%).

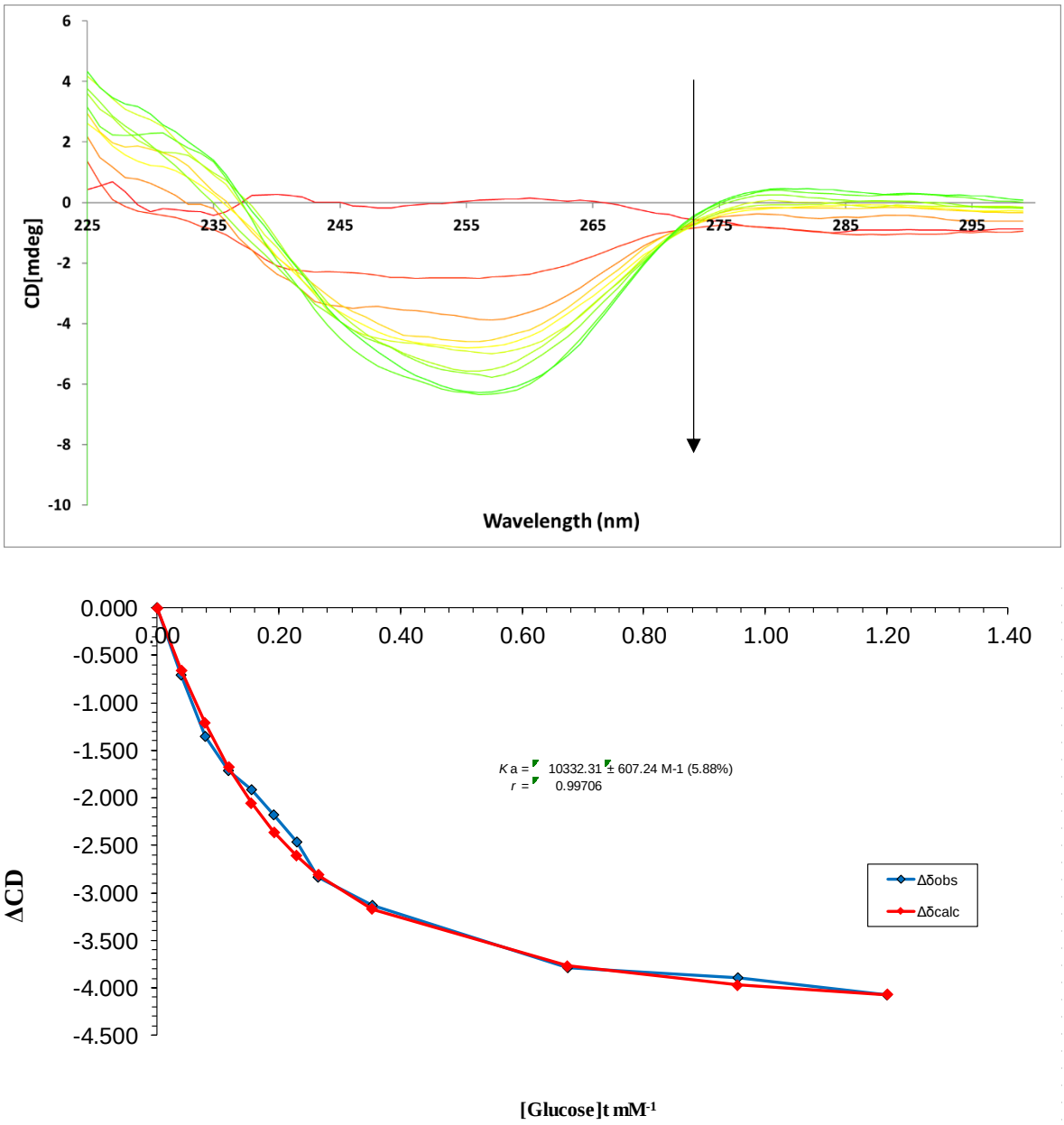


Figure 192 CD spectra (top) and binding analysis curve (bottom) for receptor **3** (0.15 mM) titrated with a combined solution of D-glucose **4** (5 mM) and receptor **3** (0.15 mM), in human serum (90% v/v, D-glucose removed) with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing guest concentration (mM). The calculated values for the ΔCD are overlaid with the observed values giving K_a = 10,300 ± 610 M^{-1} (5.9%).

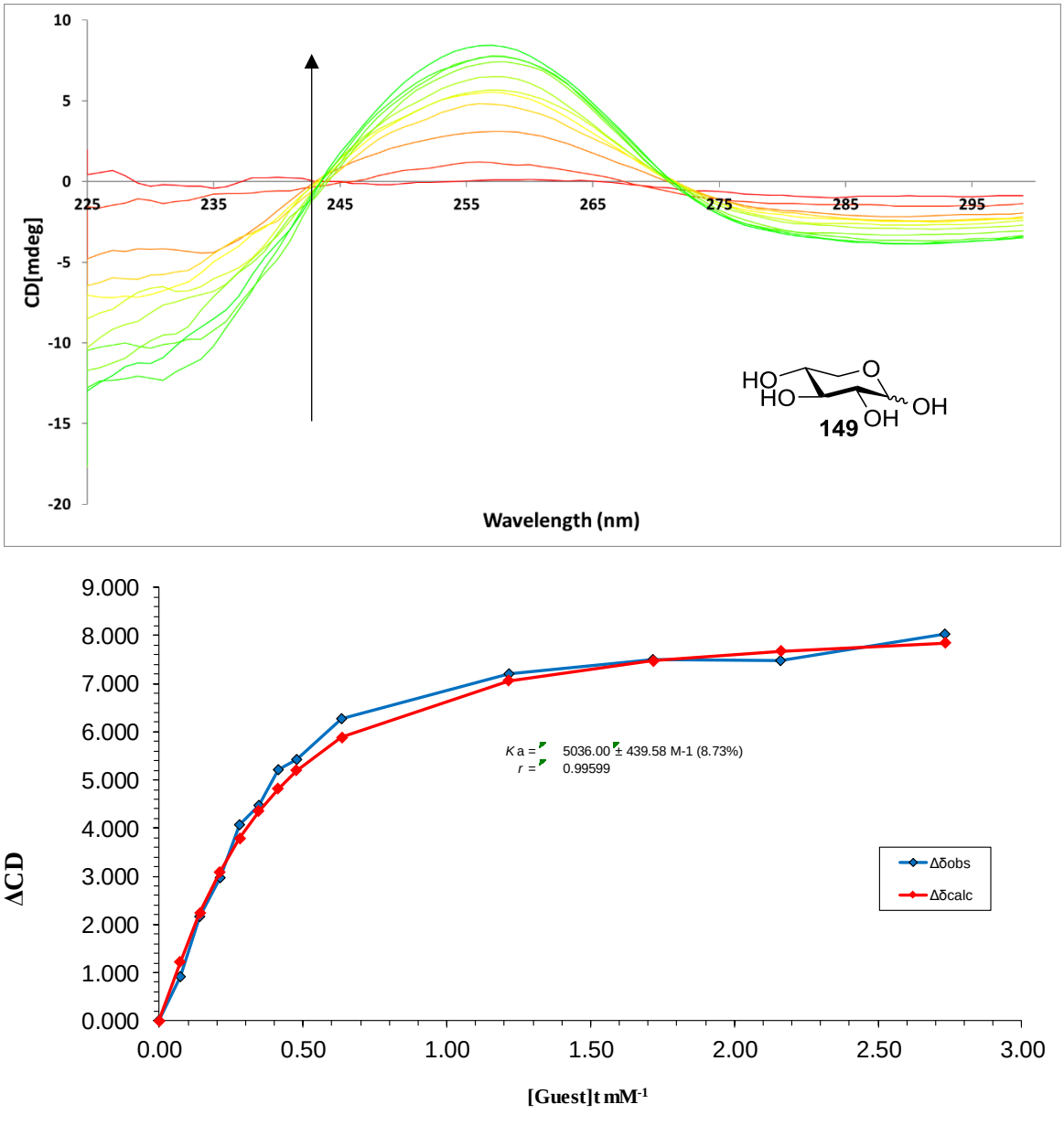


Figure 193 CD spectra (top) and binding analysis curve (bottom) for receptor **3** (0.25 mM) titrated with a combined solution of D-xylose **149** (9 mM) and receptor **3** (0.25 mM), in water with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing guest concentration (mM). The calculated values for the ΔCD are overlaid with the observed values giving K_a = 5000 ± 440 M^{-1} (8.7%).

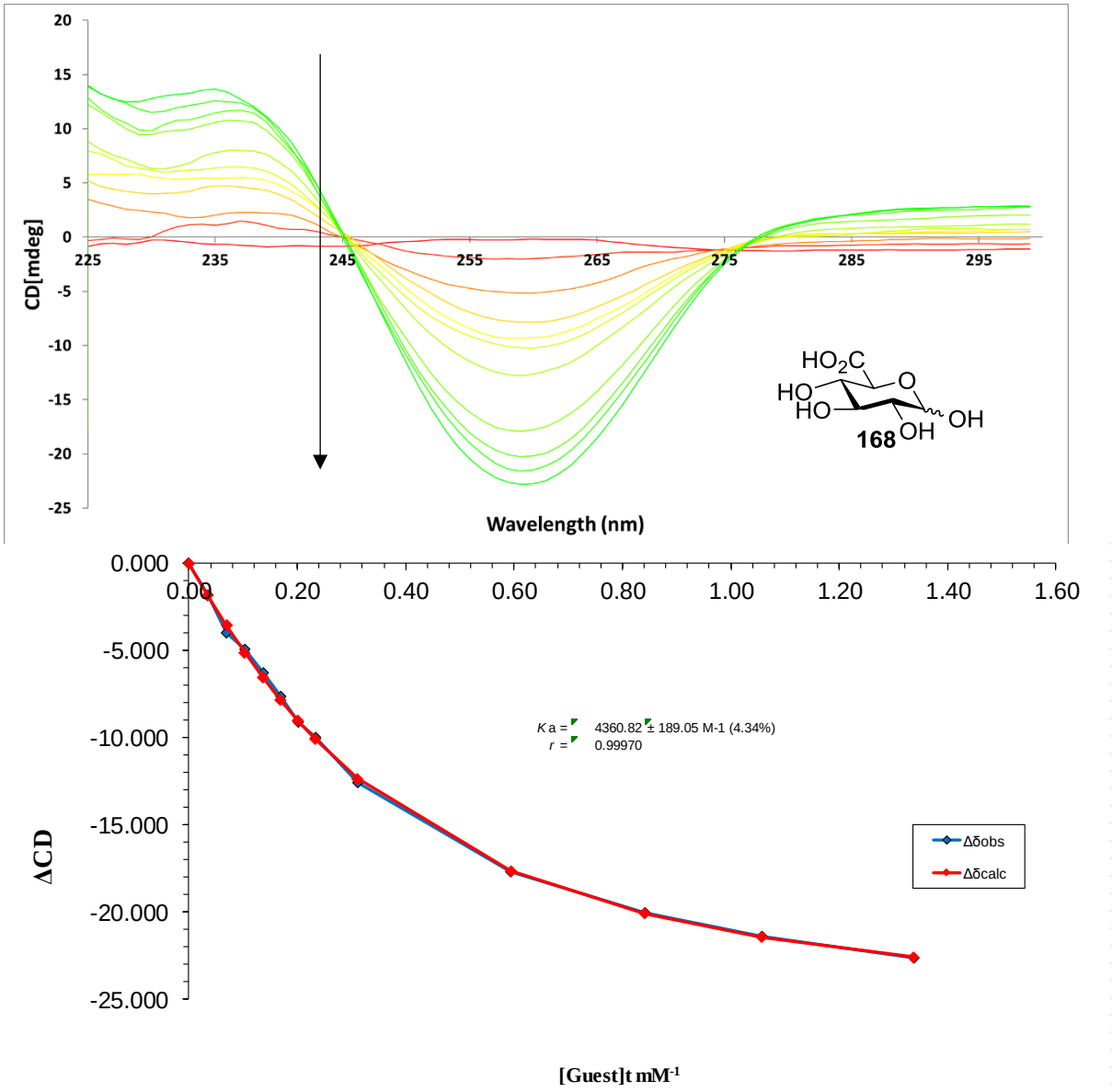


Figure 194 CD spectra (top) and binding analysis curve (bottom) for receptor **3** (0.25 mM) titrated with a combined solution of D-glucuronic acid **168** (5 mM) and receptor **3** (0.25 mM), in water with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing guest concentration (mM). The calculated values for the ΔCD are overlaid with the observed values giving K_a = 4400 ± 190 M^{-1} (4.3%).

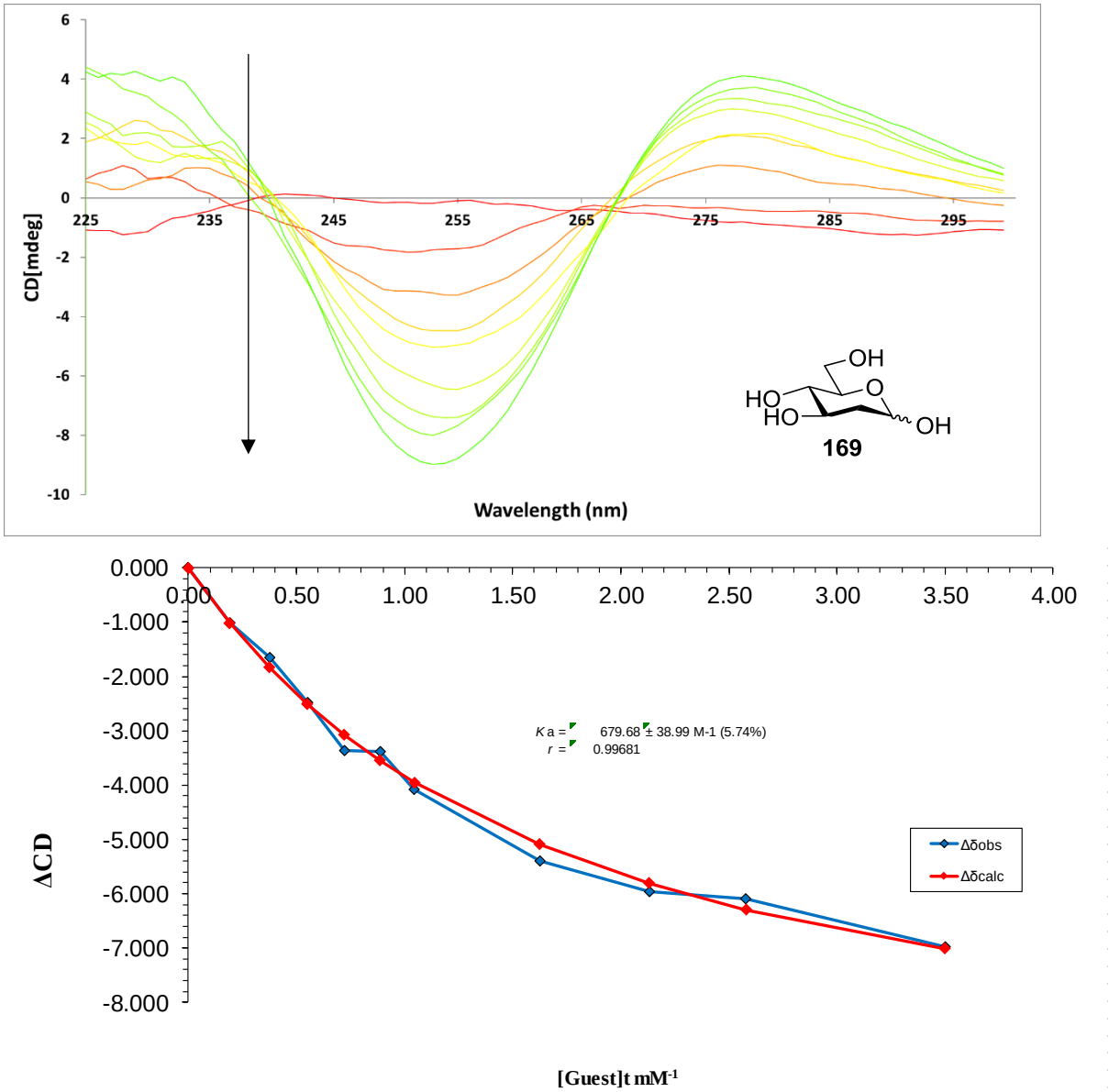


Figure 195 CD spectra (top) and binding analysis curve (bottom) for receptor **3** (0.25 mM) titrated with a combined solution of 2-deoxy-D-glucose **169** (10 mM) and receptor **3** (0.25 mM), in water with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing guest concentration (mM). The calculated values for the ΔCD are overlaid with the observed values giving K_a = 680 ± 39 M^{-1} (5.7%).

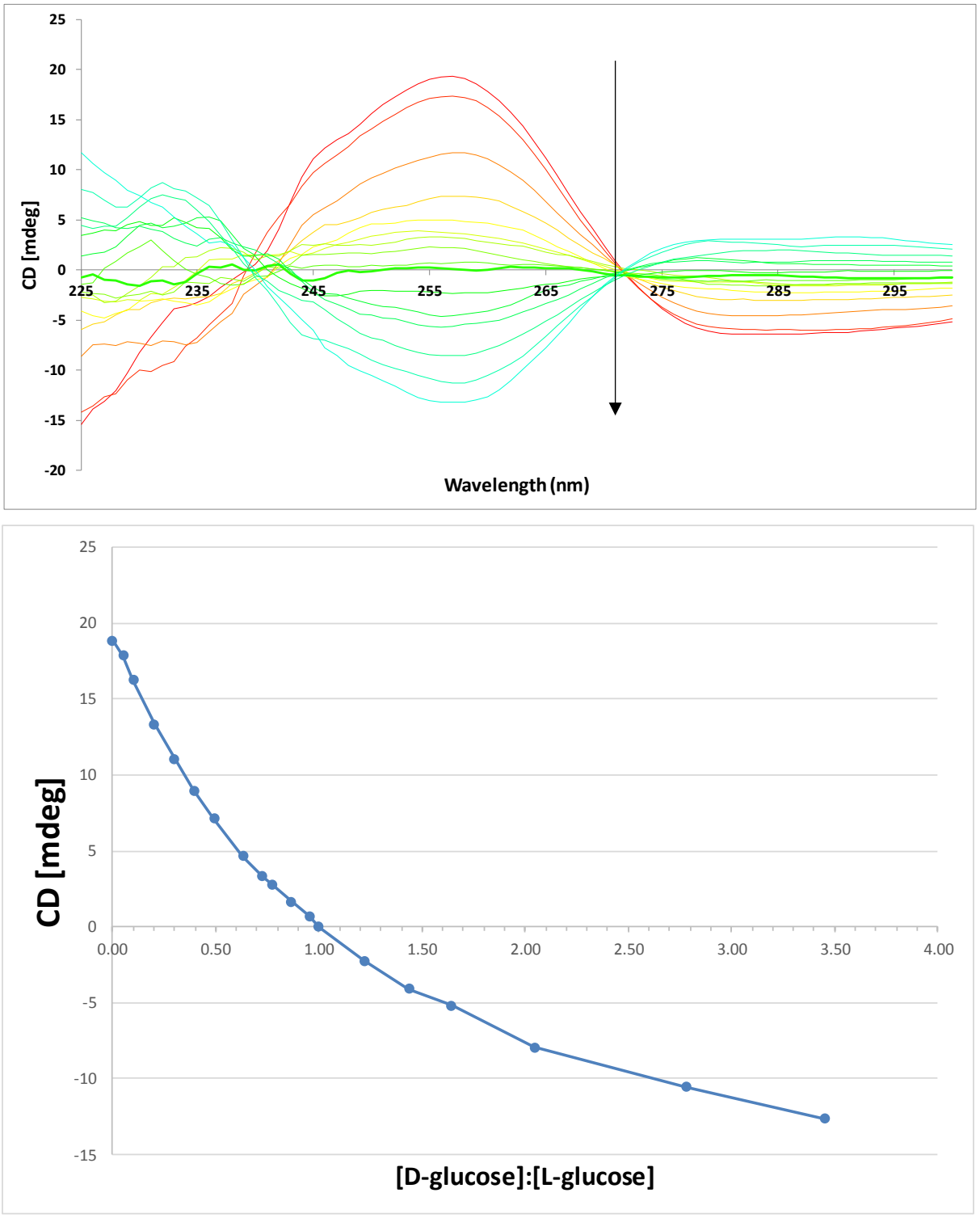


Figure 196 CD spectra (top) and analysis curve (bottom) for receptor **3** (0.25 mM) saturated with L-glucose **187** (2.4 mM) titrated with a combined solution of D-glucose **187** (10 mM), L-glucose **187** (2.4 mM) and receptor **3** (0.25 mM), in water with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing ratio of [D-glucose]:[L-glucose], with zero signal observed when ratio is equal to one.

Table 12 Raw data obtained from experiment outlined in Figure 196, where CD signal is equal to zero when ratio of D-glucose **4** and L-glucose **187** concentrations are equal (highlighted in yellow).

[D-Glc] / mM	CD[mdeg]	[DGlc]:[LGlc]
0.000	18.88	0.00
0.121	17.847	0.05
0.242	16.2479	0.10
0.480	13.3685	0.20
0.716	11.0832	0.30
0.950	8.96877	0.40
1.180	7.12859	0.49
1.520	4.65307	0.63
1.744	3.30714	0.73
1.855	2.77992	0.77
2.075	1.67586	0.86
2.292	0.707081	0.96
2.400	-0.01781	1.00
2.930	-2.23679	1.22
3.444	-4.07233	1.44
3.945	-5.18479	1.64
4.907	-7.93841	2.04
6.684	-10.5528	2.78
8.289	-12.6356	3.45

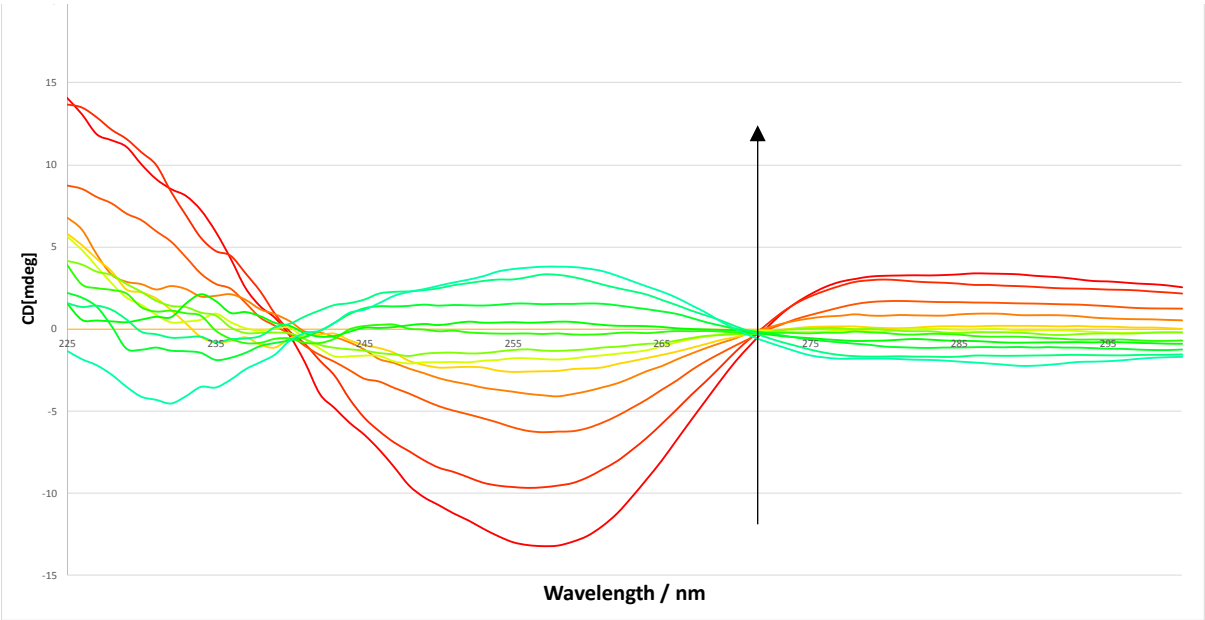


Figure 197 CD spectra for receptor **3** (0.25 mM) in human serum (50% v/v) saturated with D-glucose already present in human serum. L-glucose **187** (20 mM) and receptor **3** (0.25 mM) in human serum (50% v/v) was titrated into this medium with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing ratio of [D-glucose]:[L-glucose], with zero signal observed when ratio is equal to one. This is concentration of 50% of the D-glucose in the human serum.

Table 13 Raw data from experiment outlined in Figure 197, where CD signal is equal to zero when ratio of D-glucose **4** and L-glucose **187** concentrations are equal (highlighted in blue). Extrapolation of this data gives the glucose concentration in 50% human serum: 2.85-2.95 mM. Multiplying this by two gives the actual concentration of D-glucose in the human serum: 5.7-5.9 mM.

[D-Glc] / M	[D-Glc] / mM	CD[mdeg]	Calc
0.00	0.00	-12.63	-12.63
0.00	0.38	-9.06	-9.46
0.00	0.96	-6.06	-5.93
0.00	1.51	-3.81	-3.62
0.00	2.01	-2.42	-1.99
0.00	2.21	-1.67	-1.47
0.00	2.30	-1.24	-1.24
0.00	2.49	-0.94	-0.80
0.00	2.71	-0.34	-0.31
0.00	3.14	0.35	0.51
0.00	3.92	1.54	1.70
0.00	4.92	2.94	2.85
0.01	5.77	3.67	3.60

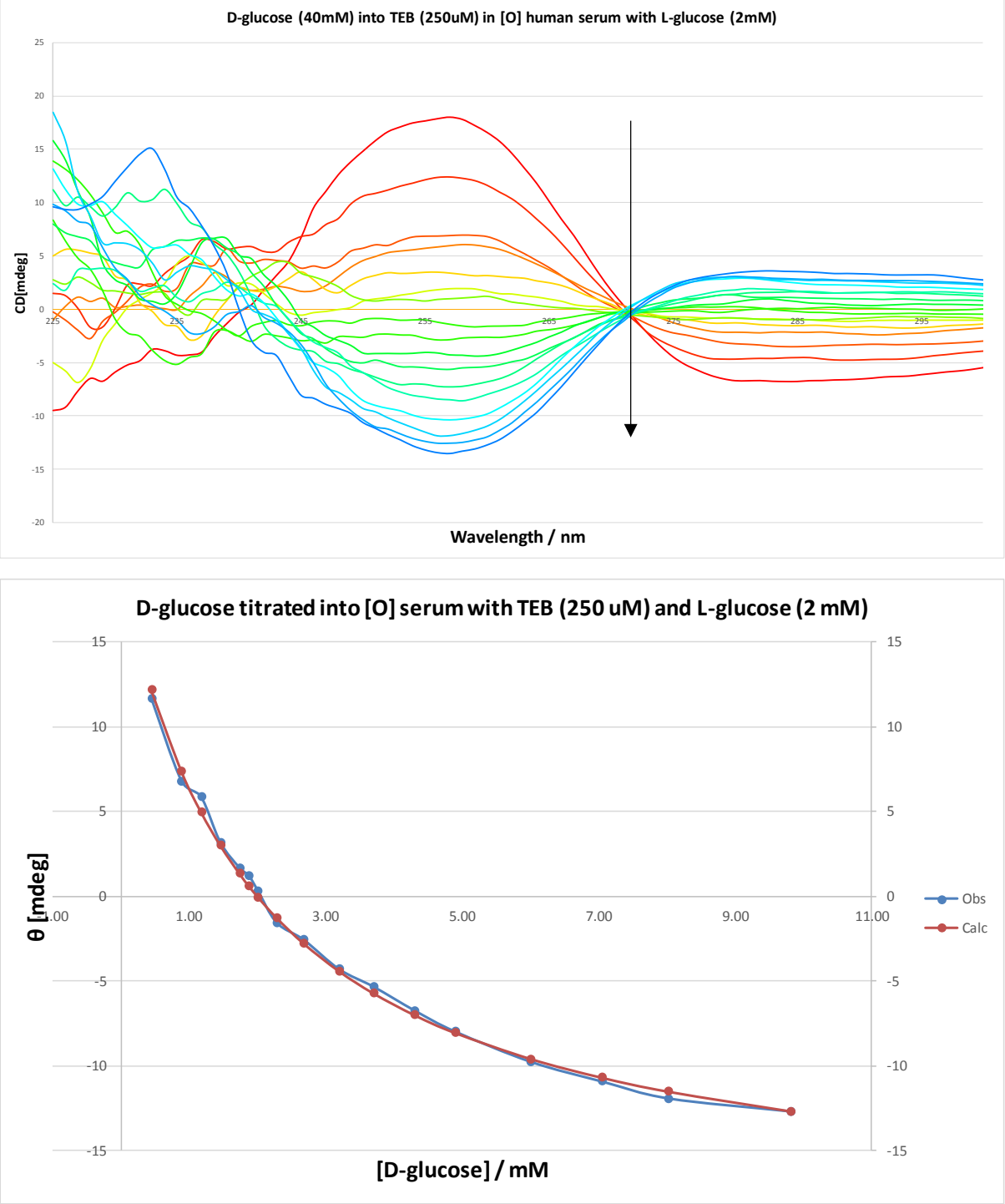
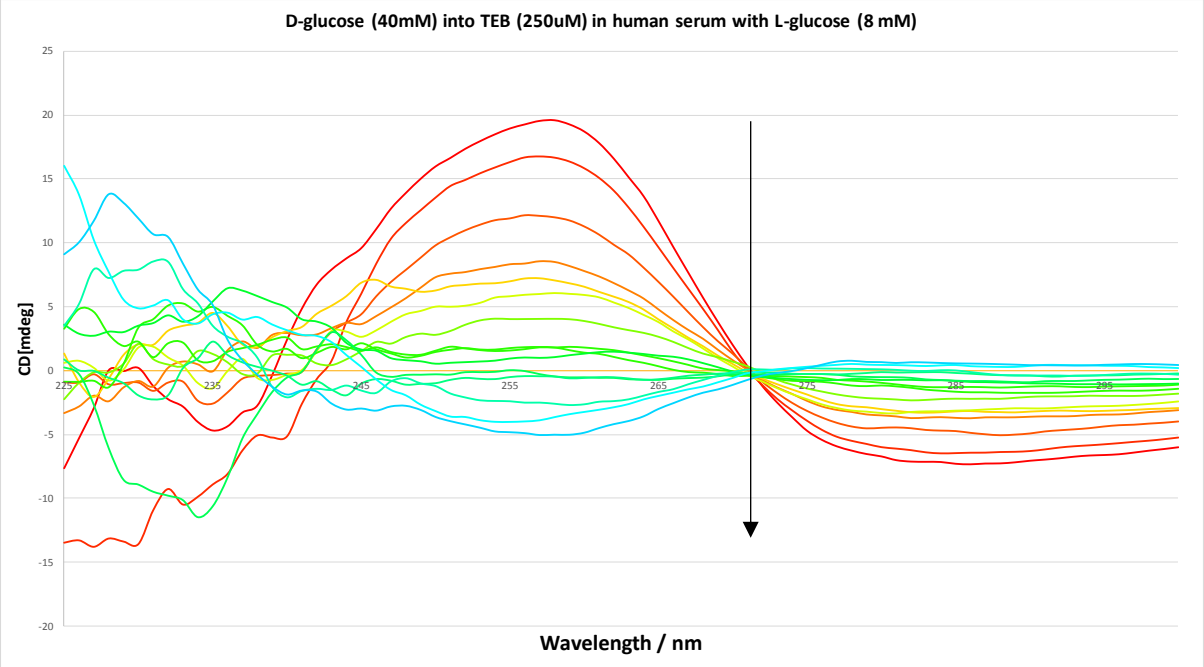


Figure 198 CD spectra (top) for receptor **3** (0.25 mM) in human serum (D-glucose removed, 50% v/v) saturated with L-glucose **187** (2 mM). D-glucose **4** (40 mM), L-glucose **187** (2 mM) and receptor **3** (0.25 mM) in human serum (D-glucose removed, 50% v/v) was titrated into this medium with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing concentration of D-glucose **4** (bottom). A calculated curve was fitted using a non linear least squares regression analysis and the resultant equation of the calculated curve was used to measure the concentration of D-glucose in a sample of the same human serum with added receptor **3** (see Table 14).

Table 14 Measured CD intensity of receptor **3** (0.25 mM) with L-glucose (2mM) in human serum (D-glucose present, 50% v/v). This CD intensity was inputted into the equation for the calculated curve outlined in Figure 198 to obtain the measured concentration of D-glucose in the human serum (50% v/v). This was repeated three times and averaged. The final averaged concentration of D-glucose was multiplied by two to obtain the actual concentration of D-glucose in the human serum (5.7 mM).

Quantification of DGlc in serum	**Obs θ**	**[D] / mM**
Run 1	-3.50	2.89
Run 2	-3.17	2.79
Run 3	-3.37	2.85
Average	-3.35	2.84
	Final [DGlc] / mM	5.69



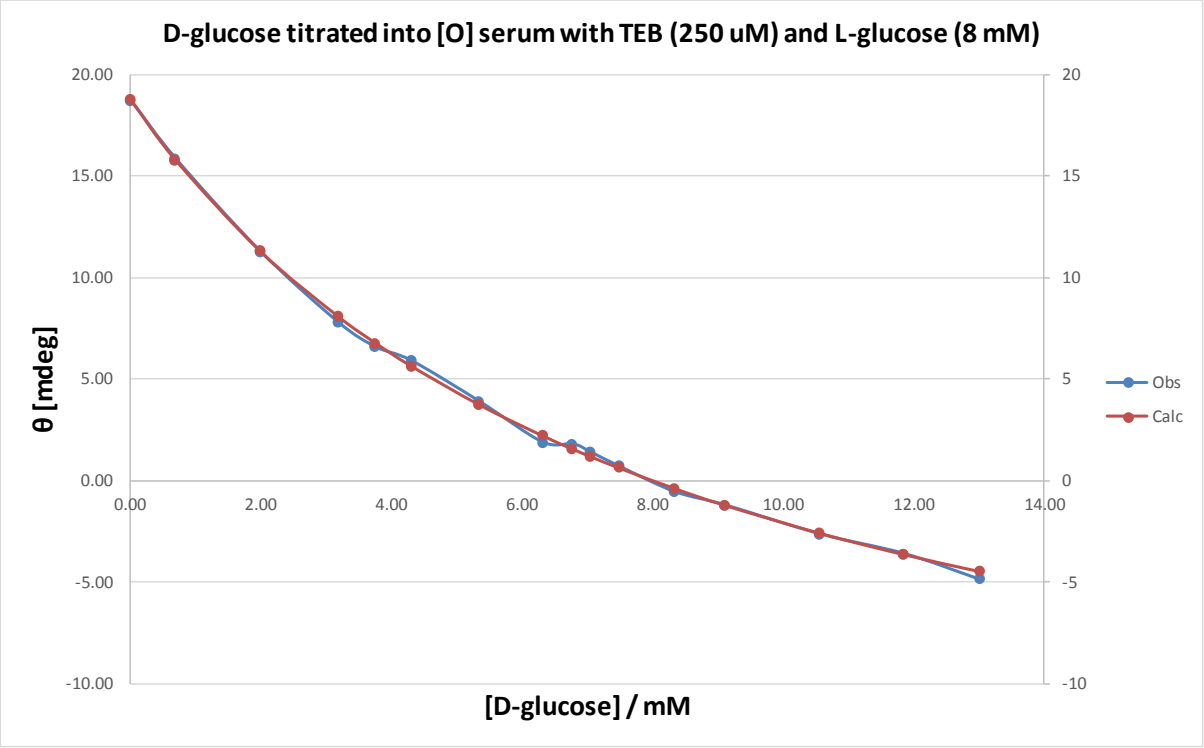


Figure 199 CD spectra (top) for receptor **3** (0.25 mM) in human serum (D-glucose removed, 50% v/v) saturated with L-glucose **187** (8 mM). D-glucose **4** (40 mM), L-glucose **187** (8 mM) and receptor **3** (0.25 mM) in human serum (D-glucose removed, 50% v/v) was titrated into this medium with 10 mM phosphate buffer (pH 7.4) at 298 K. Monitored intensity at 260 nm. Changes in CD intensity (ΔCD) were plotted against increasing concentration of D-glucose **4** (bottom). A calculated curve was fitted using a non linear least squares regression analysis and the resultant equation of the calculated curve was used to measure the concentration of D-glucose in a sample of the same human serum with added receptor **3** (see Table 15).

Table 15 Measured CD intensity of receptor **3** (0.25 mM) with L-glucose (8mM) in human serum (D-glucose present, 50% v/v). This CD intensity was inputted into the equation for the calculated curve outlined in Figure 199**Figure 198** to obtain the measured concentration of D-glucose in the human serum (50% v/v). This was repeated three times and averaged. The final averaged concentration of D-glucose was multiplied by two to obtain the actual concentration of D-glucose in the human serum (5.8 mM).

Quantification of DGlc in serum	**Obs θ**	**[D] / mM**
Run 1	8.68	2.94
Run 2	8.71	2.93
Run 3	8.93	2.84
Average	8.77	2.90
	Final [DGlc] / mM	5.80